



Towards Near-Exact Solutions of Molecular Electronic Structure: Full Coupled-Cluster Reduction with a Second-Order Perturbative Correction

Xu, Enhua

Uejima, Motoyuki

Ten-no, L. Seiichiro

(Citation)

Journal of Physical Chemistry Letters, 11(22):9775-9780

(Issue Date)

2020-11-19

(Resource Type)

journal article

(Version)

Accepted Manuscript

(Rights)

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Journal of Physical Chemistry Letters, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see <https://doi.org/10.1021/acs.jpclett.0c03084>

(URL)

<https://hdl.handle.net/20.500.14094/90007681>



Towards Near-Exact Solutions of Molecular Electronic Structure: Full Coupled-Cluster Reduction with a Second-Order Perturbative Correction

Enhua Xu,[†] Motoyuki Uejima,[†] and Seiichiro L. Ten-no^{*,†,‡}

[†]*Graduate School of Science, Technology, and Innovation, Kobe University,
Kobe 657-8501 Japan*

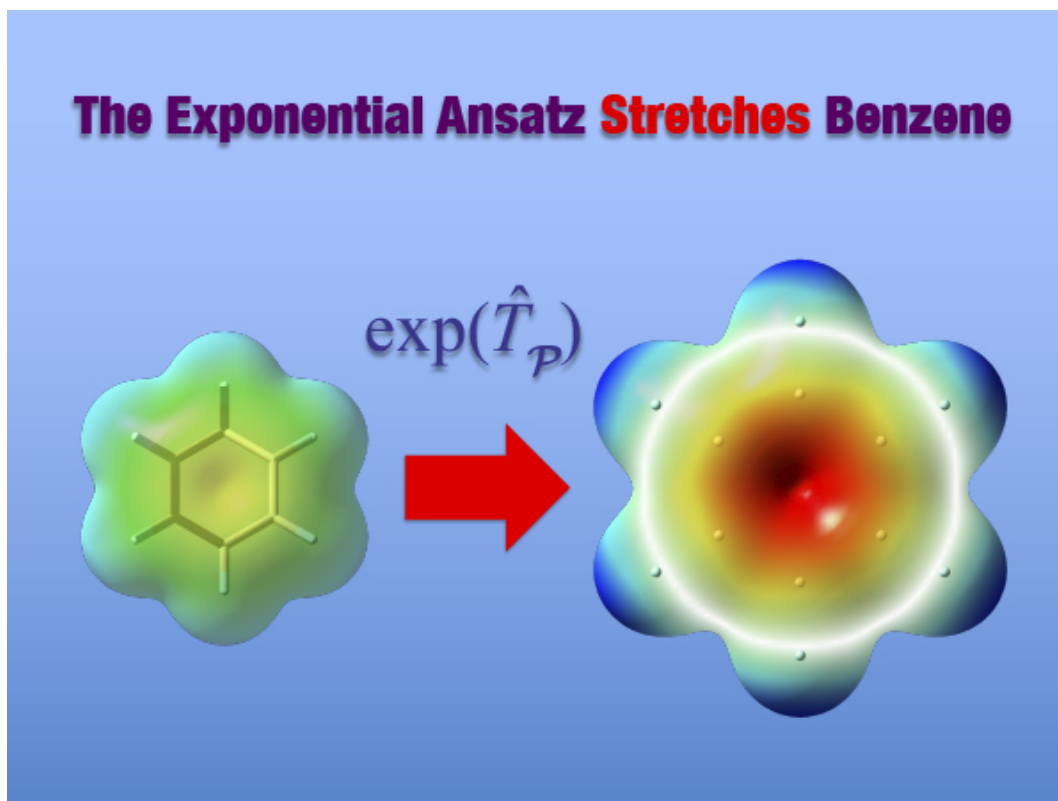
[‡]*Graduate School of System Informatics, Kobe University, Kobe 657-8501 Japan*

E-mail: tenno@garnet.kobe-u.ac.jp

Abstract

We introduce a new augmented adaptation of the recently-developed full coupled-cluster reduction (FCCR) with a second-order perturbative correction, abbreviated as FCCR(2). FCCR is a selected coupled-cluster expansion aimed at optimally reducing the excitation manifold and commutator expansions for high-rank excitations for obtaining accurate solutions of the electronic Schrödinger equation in a size-extensive manner. The present FCCR(2) enables to estimate the residual correlation of FCCR by the second-order perturbative correction $E^{(2)}$ from the complementary space of the FCCR projection manifold. The linear relationship between $E^{(2)}$ and the energy of FCCR(2) allows accurate estimates of near-exact energies for a wide variety of molecules with strong electron correlation. The potential of the method is demonstrated using challenging cases, the ground state electronic energy of the benzene molecule in equilibrium

and stretched geometries, and the isomerization energy of the transition metal complex, $[\text{Cu}(\text{NH}_3)_2\text{O}_2]^{2+}$.



Devising a reliable model for describing correlated electrons in molecules from first principles is one of the most important challenges in the quantum theory of chemistry and physics. Coupled-cluster (CC) theory^{1,2} has enabled both accuracy and efficiency owing to the exponential ansatz of the wave function to cover an encompassing range of applications. Especially, CC singles, doubles with perturbative triples (CCSD(T))³ is referred to as 'the gold standard' of *ab initio* theory, and has been widely used to treat single-reference (SR) molecules dominated by the Hartree–Fock (HF) wave functions with high accuracy. Nevertheless, the CC ansatz for strongly-correlated systems requires higher-rank cluster operators, and is usually formidable within the standard SR framework. Multi-reference (MR) extensions of CC have partially mitigated this problem,^{4,5} although the prescription of MRCC is yet to be well-established. In particular, the treatment of a large active space is a crucial issue for practical applications since the majority of MRCC employs a configuration interaction (CI) of complete active space (CAS) as a reference wave function, that is, the full CI (FCI) within the active space.

To ameliorate these features, we recently introduced a novel selected CC approach, the so-called full CC reduction (FCCR).⁶ The exact solution of the N -electronic Schrödinger equation Ψ_N can be expressed either in the linear (FCI) or exponential (FCC) form,

$$|\Psi_N\rangle = (1 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + \cdots + \hat{C}_N)|0\rangle \quad (1)$$

$$= \exp(\hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N)|0\rangle, \quad (2)$$

where, \hat{C}_k and \hat{T}_k denote the k -fold excitation operators with respect to a suitable Fermi vacuum $|0\rangle$ of a single Slater-determinant. FCCR makes the best of the sparsity of the exponential ansatz consisting of only connected operators by employing screenings to reduce the projection manifold and commutator operations for higher excitations.⁶ The exponential ansatz guarantees the proper scaling of the energy with the system size, and the resulting size-extensivity is the obvious advantage of FCCR over the recent adaptive and stochastic CI

approaches.^{7–14} The projection manifold of FCCR is iteratively updated using the interacting space connected to the primary clusters. Screening is efficiently performed by taking account of the contribution of exclusion-principle-violating (EPV) terms.⁶

In this Letter, we propose an augmented FCCR with a second-order perturbative correction, termed FCCR(2), inspired by the success of CCSD(T) in the SR case. We also present the implementation of FCCR(2) along with applications to challenging systems.

The projection manifold of FCCR with high precision contains an enormous number of high-rank cluster operators with small amplitudes, and the iterative treatment of all of them hinders the wide applicability of the method. We, therefore, divide the projection manifold and the corresponding cluster operators into those in the primary space \mathcal{P} and secondary space \mathcal{Q} . The FCCR wave function is parametrized within the manifold \mathcal{P} ,

$$|\Psi_{\text{FCCR}}\rangle = e^{\hat{T}_{\mathcal{P}}} |0\rangle, \quad (3)$$

where $\hat{T}_{\mathcal{P}} = \sum_{\mu \in \mathcal{P}} t_{\mu} \hat{a}_{\mu}^{\dagger}$, \hat{a}_{μ}^{\dagger} denotes the excitation operator for the basis $|\mu\rangle$: $\hat{a}_{\mu}^{\dagger} |0\rangle = |\mu\rangle$. We employ the HF wave function for $|0\rangle$ throughout this work. The amplitudes $\{t_{\mu}\}$ and FCCR energy are obtained by solving the working equation.⁶ $\hat{T}_{\mathcal{P}}$ is comprised of cluster operators with large amplitudes to describe static and the majority of dynamic correlation effects. Then, the remaining correlation energy from \mathcal{Q} is calculated from a second-order perturbative correction $E^{(2)}$ for the FCCR(2) energy

$$E_{\text{FCCR}(2)} = E_{\text{FCCR}} + E^{(2)}. \quad (4)$$

The survey of FCCR(2) is depicted in FIG. 1.

The primary space \mathcal{P} is iteratively updated with the aid of the generator space \mathcal{G} as a subset of \mathcal{P} . The present implementation of FCCR employs two screening parameters for the projection manifold, *i.e.*, the principal screening threshold $\vartheta_{\mathcal{P}}$ to select the cluster operators for \mathcal{P} and generator threshold $\vartheta_{\mathcal{G}}$ for constructing \mathcal{G} . The update procedure is given

as follows.

Update of \mathcal{P} :

Initialize the primary and generator spaces, \mathcal{P}_0 and \mathcal{G}_0 to be $|0\rangle$, and $n = 1$.

Step 1 ($\mathcal{G}_{n-1} \rightarrow \mathcal{P}_n$): All single and double (SD) excitations with respect to \mathcal{G}_{n-1} are generated, and those fulfilling $\left| \frac{\sigma_\kappa}{\Delta_\kappa} \right| \geq \vartheta_{\mathcal{P}}$ are added to \mathcal{P}_{n-1} to form \mathcal{P}_n , using the objects σ_κ and Δ_κ in Eqs. (7) and (8).

Step 2 ($\mathcal{P}_n \rightarrow E^{(n)}$): Solve the FCCR working equation in \mathcal{P}_n for $E^{(n)}$. If $|E^{(n)} - E^{(n-2)}| < 0.5 \text{ mH}$ and $|E^{(n)} - E^{(n-1)}| < 0.1 \text{ mH}$, the iteration for \mathcal{P} is terminated.

Step 3 ($\mathcal{P}_n \rightarrow \mathcal{G}_n$): \mathcal{G}_n is updated by adding the determinants in \mathcal{P}_n with $|t_\mu| > \vartheta_{\mathcal{G}}$. Increase $n = n + 1$ and return to Step 1. $\vartheta_{\mathcal{G}}$ is fixed to be 0.01 in this work.

In the above algorithm, \mathcal{G} plays the role of spanning a reference space that can accommodate the excitations necessary to capture strong or non-dynamic correlation in addition to the Fermi vacuum $|0\rangle$. FCCR reduces to CCSD when \mathcal{G} contains only $|0\rangle$ and $\vartheta_{\mathcal{P}} = 0$. However, the calculation with $\vartheta_{\mathcal{P}} = 0$ becomes impractical for large molecules even in the SR case. Alternatively, we use finite $\vartheta_{\mathcal{P}}$ in conjunction with the extrapolation of the perturbative correction. The distribution of the excitation levels of \mathcal{P} or \mathcal{G} is not always continuous, especially for strongly correlated systems, *e.g.*, some doubles and quadruples are strongly coupled via two-electron interactions although they are not connected by single excitations. Therefore, we update the projection manifold using the SD space from \mathcal{G} to construct a compact \mathcal{P} with discrete excitation levels. The above procedure is quite efficient and reaches the convergence in several iterations. Operations with negligible contributions arising from the nonlinear structure of the CC working equation are discarded using the operation screening threshold $\vartheta_{\mathcal{O}}$ in conjunction with the in the EPV form of the equation.⁶ A large operation threshold $\vartheta_{\mathcal{O}}$ may be employed to update \mathcal{P} .

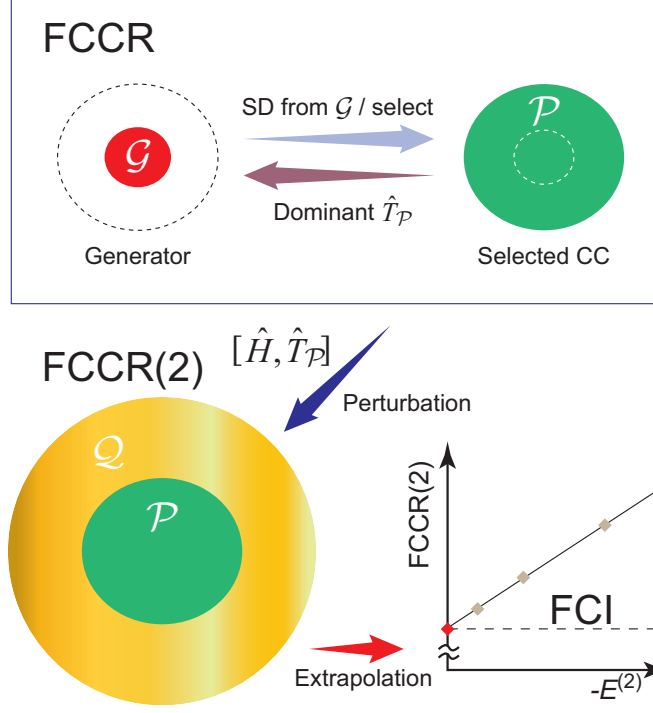


Figure 1: Survey of the FCCR(2) scheme. The primary space \mathcal{P} is iteratively updated with the aid of the the generator space \mathcal{G} in combination with the FCCR calculation. The residual correlation energy from the secondary space \mathcal{Q} is estimated by second-order correction.

Once we obtain the amplitudes $\{\hat{T}_\mu\}$ and E_{FCCR} , the secondary space \mathcal{Q} is generated for $E^{(2)}$ using the interaction space connected through the single commutator $[\hat{H}, \hat{T}_\mathcal{P}]$.⁶ The simplest example is all triples generated from $\hat{T}_\mathcal{P}$ spanning CCSD. In parallel to the second-order perturbative corrections in SRCC,^{15,16} we employ formulas based on the order-by-order expansion in the bi-orthogonal basis for the similarity transformed Hamiltonian, $\bar{H}_\mathcal{P} = e^{-\hat{T}_\mathcal{P}} \hat{H} e^{\hat{T}_\mathcal{P}}$. The second-order correction is expressed as

$$E^{(2)} = \sum_{\kappa \in \mathcal{Q}} \frac{\eta_\kappa \sigma_\kappa}{\Delta_\kappa}, \quad (5)$$

where the objects are

$$\eta_\kappa = \langle 0 | (1 + \hat{\Lambda}_\mathcal{P}) \bar{H}_\mathcal{P} | \kappa \rangle, \quad (6)$$

$$\sigma_\kappa = \langle \kappa | \bar{H}_{\mathcal{P}} | 0 \rangle, \quad (7)$$

$\hat{\Lambda}_{\mathcal{P}}$ is the left-hand eigen state vector of $\bar{H}_{\mathcal{P}}$: $\hat{\Lambda}_{\mathcal{P}} = \sum_{\mu \in \mathcal{P}} l_\mu \hat{a}_\mu$ with the amplitudes $\{l_\mu\}$. We use the simple denominator in the Møller-Plesset form¹⁶ in this particular work,

$$\Delta_\kappa = -\langle \kappa | [\hat{F}, \hat{a}_\kappa^\dagger] | 0 \rangle, \quad (8)$$

with the single-electron Fock operator \hat{F} from the HF equation. In addition, we always monitor the second order correction $E^{(\text{EN}2)}$ using the Epstein-Nesbet (EN) type denominator

$$\Delta_\kappa^{(\text{EN})} = E_{\text{FCCR}} - \langle \kappa | \hat{H} | \kappa \rangle, \quad (9)$$

instead of Δ_κ to check the uncertainty from the partitioning. The computation of the objects, $\{\eta_\kappa\}$ and $\{\sigma_\kappa\}$, proportional to the dimension of \mathcal{Q} is considerably more expensive than that of $\hat{T}_{\mathcal{P}}$ and $\hat{\Lambda}_{\mathcal{P}}$. The screenings for the nonlinear operations with the EPV form of the working equations are generalized for these objects by excluding all $\hat{T}_{\mathcal{P}}$ and $\hat{\Lambda}_{\mathcal{P}}$ amplitudes relevant to EPV from the screening. It is found that $E^{(2)}$ is relatively insensitive to the operation threshold $\vartheta_{\mathcal{O}}$ compared to E_{FCCR} , and a larger $\vartheta_{\mathcal{O}}$ can be employed for $E^{(2)}$. The explicit forms of the working equations for $\{l_\mu\}$, $\{\sigma_\kappa\}$ and $\{\eta_\kappa\}$ are detailed in the Supporting Information. On practicing the second-order correction, the efficiency is increased by excluding insignificant terms of \mathcal{Q} contributions to $E^{(2)}$. $\{\eta_\kappa\}$ containing \hat{a}_κ^\dagger can be generated much faster than $\{\sigma_\kappa\}$. And thus, we calculate $\{\eta_\kappa\}$ first and then exclude them satisfying $\left| \frac{\eta_\kappa}{\Delta_\kappa} \right| < \vartheta_{\mathcal{F}}$ from the summation of Eq. (5). $\vartheta_{\mathcal{F}}$ is fixed to be 2×10^{-7} in this work. These objects are calculated by distributing the loop over κ without a large memory requirement, as \mathcal{Q} is divided into batches. All FCCR(2) calculations are performed using the GELLAN program¹⁷ in a massively parallel implementation.

Recently, extrapolation schemes for selected CI approaches with perturbative corrections have been proposed.^{18–21} Despite $E^{(2)}$ for FCCR that is radically different from those for CI

in the structure and the complexity of the required objects, a linear relationship between $E_{\text{FCCR}(2)}$ and $E^{(2)}$ also holds for FCCR(2). This enables accurate extrapolations of FCCR(2) to near FCI limits with smaller extrapolation distances.

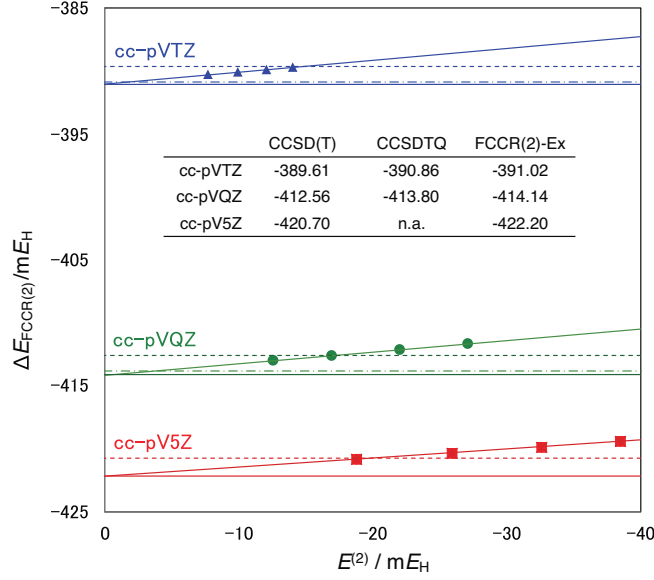


Figure 2: FCCR(2) correlation energies for N_2 in the cc-pVXZ basis set (X=T, Q, 5) basis sets at $R_{\text{NN}}=2.068$ a.u. using different $\vartheta_{\mathcal{P}} = 5 \times 10^{-4}$, 4×10^{-4} , 3×10^{-4} and 2×10^{-4} . The operation thresholds are fixed as $\vartheta_{\mathcal{O}} = 10^{-4}$, 10^{-7} , and 3×10^{-6} for the update of \mathcal{P} , E_{FCCR} , and $E^{(2)}$, respectively. The horizontal lines denote CCSD(T) (dashed), CCSDTQ (dash-dotted), and FCCR(2)-Ex (solid) results.

We first present the performance of FCCR(2) for the N_2 molecule near the equilibrium bond distance in cc-pVXZ (X=T, Q, 5) basis sets.²² FIG. 2 shows the correlation energy of FCCR(2) as a function of $E^{(2)}$ by changing $\vartheta_{\mathcal{P}}$ in comparison with CCSD(T) and CC singles, doubles, triples, and quadruples (CCSDTQ).²³ $\Delta E_{\text{FCCR}(2)}$ shows nearly perfect linear correlation to $E^{(2)}$ with the root mean square deviations, $\Delta E_{\text{RMSD}} = 0.007$, 0.017 , and 0.027 mE_{H} for cc-pVTZ, QZ, and 5Z, respectively. Consequently, we obtain accurate extrapolations of FCCR(2) by considering the limit of $E^{(2)} = 0$, denoted by FCCR(2)-Ex. The energy differences of CCSD(T), CCSDTQ, and FCCR(2)-Ex are nearly independent of the cardinal number of the basis set, indicating that the majority of the residual correlation

beyond CCSD(T) is not short-ranged, such as the electron-electron cusps that can be efficiently described by the F12 ansatz.^{24–28} Note that the FCCR excitation manifold is rather different from the standard CC, *e.g.* the generator space \mathcal{G} includes 177 singles and doubles to induce a total 40,919 determinants in \mathcal{P} space up to quadruples in the cc-pV5Z case.

We apply FCCR(2) to the benzene molecule in the cc-pVDZ basis set.²² This system was employed in a recent blind test at the equilibrium geometry²⁹ for the frozen-core ground state energy of highly-accurate electronic structure methods including CCSDTQ, FCCR(2), many-body expansion FCI (MBE-FCI),^{30,31} selected CI with perturbative corrections,^{18–21} density-matrix renormalization group approach (DMRG),³² adaptive-shift full CI quantum Monte Carlo (AS-FCIQMC),³³ and cluster-analysis-driven FCIQMC (CAD-FCIQMC),³⁴ followed by more recent reports using phase-less auxiliary-field QMC (ph-AFQMC)³⁵ and CI of a perturbative selection made iteratively (CIPSI).³⁶ The total correlation energy in the equilibrium geometry was estimated to be approximately $-863 \text{ m}E_{\text{H}}$. Due to the SR nature of the system, CCSDTQ gives a rather accurate energy, $\Delta E_{\text{CCSDTQ}} = -862.37 \text{ m}E_{\text{H}}$. The FCCR(2)-Ex energy amounts to $\Delta E_{\text{FCCR(2)-Ex}} = -862.83 \text{ m}E_{\text{H}}$, which is considered as one of the most accurate results. We elaborate on the calculation of FCCR(2). Nevertheless, a more challenging system is stretched benzene. Thus, we also examine benzene in a stretched geometry, where all bond distances are extended 1.5 times of the equilibrium geometry. The screening thresholds are the same as those used for N_2 .

Table 1 lists the dimensions of the spaces of \mathcal{G} , \mathcal{P} , and \mathcal{Q} in the spin-orbital basis; $\Delta E^{\text{FCCR(2)}}$ and $E^{(2)}$ with the tightest principal threshold; $\vartheta_{\mathcal{P}} = 2 \times 10^{-4}$ along with $\Delta E_{\text{FCCR(2)-Ex}}$, extrapolation distance, ΔE_{dist} , and ΔE_{RMSD} from the extrapolations in the equilibrium and stretched geometries. In the equilibrium geometry, \mathcal{P} comprises of excitations from singles to quadruples, and all FCCR amplitudes are less than 0.025. In contrast, the stretched benzene is a strongly correlated MR system, and the distribution of the excitation manifold of \mathcal{P} is extended up to septuple excitation levels with 34 amplitudes exceeding 0.1. Owing to the efficient exponential ansatz, $N_{\mathcal{P}}$ is only doubled from the equilibrium geometry, although

Table 1: Dimensions of the \mathcal{G} , \mathcal{P} , and \mathcal{Q} spaces; $\Delta E^{\text{FCCR}(2)}$, and $E^{(2)}$ for the FCCR(2) calculations with $\vartheta_{\mathcal{P}} = 2 \times 10^{-4}$; the extrapolated energies, $\Delta E_{\text{FCCR}(2)\text{-Ex}}$ and $\Delta E_{\text{FCCR}(\text{EN}2)\text{-Ex}}$, distances ΔE_{dist} , and RMSD, ΔE_{RMSD} . The energies are in $\text{m}E_{\text{H}}$.

	C_6H_6	
	equilibrium	stretched
$N_{\mathcal{G}}$	116	1,273
$N_{\mathcal{P}}$	229,842	411,628
$N_{\mathcal{Q}}$	6.4×10^8	5.9×10^9
$\Delta E_{\text{FCCR}(2)}$	-856.83	-1,238.63
$E^{(2)}$	-35.10	-73.55
$\Delta E_{\text{FCCR}(2)\text{-Ex}}$	-862.82	-1,252.70
$\Delta E_{\text{FCCR}(\text{EN}2)\text{-Ex}}$	-863.67	-1,253.69
ΔE_{dist}	-5.99	-14.07
$(\%)^a$	0.69	1.12
ΔE_{RMSD}	0.03	0.02

^aPercentage of ΔE_{dist} in $\Delta E_{\text{FCCR}(2)\text{-Ex}}$

the dimension of \mathcal{Q} for $E^{(2)}$ is increased by one order of magnitude. The computational cost of FCCR(2) roughly scales as $N_{\mathcal{P}}N_{\mathcal{Q}}$. $N_{\mathcal{P}}$ are only 0.04 % and 0.01 % of the corresponding $N_{\mathcal{Q}}$ for the equilibrium and stretched benzene, respectively to recover approximately 95% of the near-exact correlation energy of $\Delta E_{\text{FCCR}(2)\text{-Ex}}$ using the small excitation manifolds. FCCR(2) recovers approximately 99% of the total correlation energy, and the remaining *ca.* 1% is compensated by the extrapolation. The deviation from $\Delta E_{\text{FCCR}(\text{EN}2)\text{-Ex}}$ (0.85 and 0.99 $\text{m}E_{\text{H}}$ for equilibrium and stretched) is less than 0.1% of the total correlation energy. It appears that the powerful exponential ansatz of $|\Psi_{\text{FCCR}}\rangle$ is effective both for accurate perturbative corrections and extrapolations on top of FCCR in comparison to the selected CI strategies. We consider the FCCR(2)-Ex energy for the stretched benzene to have similar accuracy as that in the equilibrium geometry. The types of correlated methods that can approach the FCI solution are open for future researches.

Finally, we present application of our method to a transition metal complex, $[\text{Cu}(\text{NH}_3)]_2\text{O}_2^{2+}$. It is a well-known bonding motif in biologically relevant molecules related to the peroxo and bis(μ -oxo) isomers of a $\text{Cu}_2\text{O}_2^{2+}$ core corresponding to the II and III oxidation states of the copper atoms. The interconversion energy between the two isomers of the $[\text{Cu}(\text{NH}_3)]_2\text{O}_2^{2+}$

model system has been extensively studied using the completely renormalized CC (CR-CC) methods,³⁷ restricted active space second-order perturbation theory (RASPT2),³⁸ CAS second-order perturbation theory combined with DMRG (DMRG-CASPT2) with scalar relativistic effect,³⁹ and more recent FCIQMC with strongly contracted N -electron valence second-order perturbation theory (FCIQMC-sc-NEVPT2).⁴⁰ Although most studies predict that the peroxo isomer is more stable than the bis(μ -oxo), there remain large inconsistencies in the interconversion energy,

$$\epsilon = E_{\text{bis}(\mu\text{-oxo})} - E_{\text{peroxo}}, \quad (10)$$

i.e., $\epsilon_{\text{RASPT2}}=25.2$,³⁸ $\epsilon_{\text{DMRG-CASPT2}}=22.6$,³⁹ and $\epsilon_{\text{FCIQMC-sc-NEVPT2}}=60.0\text{kcal/mol}$.⁴⁰ RASPT2 employed an active space consisting of 24 electrons in 28 orbitals (24,28), and those in DMRG-CASPT2 and FCIQMC-sc-NEVPT2 are (24,24). We perform FCCR(2) calculations using the same geometries,³⁹ and cc-pVTZ basis set for Cu and cc-pVDZ for others keeping the core orbitals frozen. The screening parameters are identical to those in the previous cases except for $\vartheta_{\mathcal{O}}$, which is 3×10^{-7} for FCCR and 3×10^{-5} for the perturbative correction to provide a reliable energy difference between the isomers.

Table 2: Dimensions of the spaces and energies of FCCR(2) and statistical measures for FCCR(2)-Ex for the peroxo and bis(μ -oxo) isomers of $[\text{Cu}(\text{NH}_3)]_2\text{O}_2^{2+}$.

	$[\text{Cu}(\text{NH}_3)]_2\text{O}_2^{2+}$	
	peroxo	bis(μ -oxo)
$N_{\mathcal{G}}$	323	407
$N_{\mathcal{P}}$	625,816	643,310
$N_{\mathcal{Q}}$	3.5×10^{10}	4.4×10^{10}
$\Delta E_{\text{FCCR}(2)}$	-1,820.63	-1,882.74
$E^{(2)}$	-138.23	-146.62
$\Delta E_{\text{FCCR}(2)\text{-Ex}}$	-1,832.90	-1,897.44
$\Delta E_{\text{FCCR}(\text{EN}2)\text{-Ex}}$	-1,829.94	-1,895.28
ΔE_{dist}	-12.27	-14.70
(%) ^a	0.67	0.77
ΔE_{RMSD}	0.16	0.10
^a Percentage of ΔE_{dist} in $\Delta E_{\text{FCCR}(2)\text{-Ex}}$		

Table 2 shows the details of spaces and energies for the isomers. The dimensions of \mathcal{G} indicate that both the peroxo and bis(μ -oxo) complexes are not as strongly correlated as the stretched benzene, although they require much larger working spaces \mathcal{Q} for perturbative corrections, indicating the presence of high-rank cluster operators in the FCCR calculations. Indeed, the maximum excitation levels of FCCR are quadruples and quintuples for the peroxo and bis(μ -oxo) isomers, respectively. The final energies of $\Delta E_{\text{FCCR}(2)\text{-Ex}}$ for the two isomers are considered to be reliable from the relatively small measures for the extrapolations, ΔE_{dist} and ΔE_{RMSD} , although the deviation from $\Delta E_{\text{FCCR}(\text{EN}2)\text{-Ex}}$ is somewhat increased compared to the benzene case.

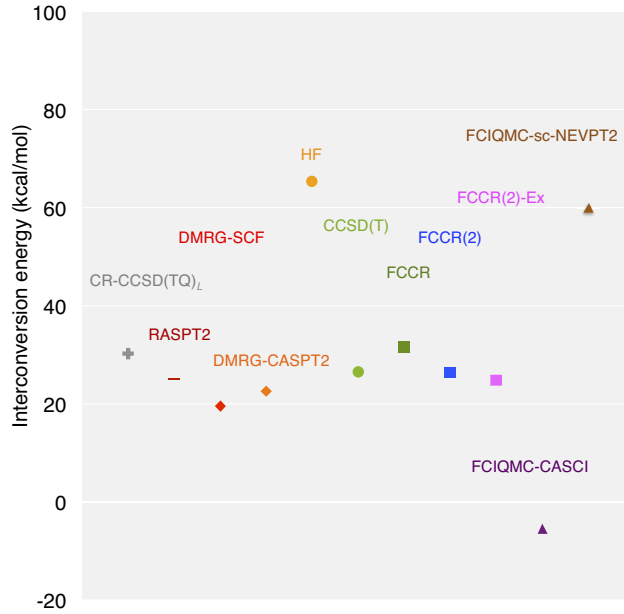


Figure 3: Interconversion energy (kcal/mol) between the peroxo and bis(μ -oxo) isomers of $[\text{Cu}(\text{NH}_3)]_2\text{O}_2^{2+}$ of various electronic structure calculations.

FIG. 3 summarizes the interconversion energies of the isomers from various electronic structure calculations. The FCCR interconversion energy is $\epsilon_{\text{FCCR}} = 31.65$ kcal/mol. The second-order perturbative correction reduces this by 5 kcal/mol to yield $\epsilon_{\text{FCCR}(2)} = 26.38$

kcal/mol. Our best estimate of the extrapolated FCCR(2) is $\epsilon_{\text{FCCR}(2)\text{-Ex}} = 24.86$ kcal/mol in quantitative agreement with RASPT2 employing the relatively large active space (24,28). Although the FCI limit in the basis set of RASPT2 is not available at this stage, CR-CCSD(TQ)_L also appears to provide accurate result, 30.2 kcal/mol³⁷ in the same basis set. CCSD(T) gives a reasonable result $\epsilon_{\text{CCSD(T)}} = 26.6$ kcal/mol under the same conditions of FCCR, presumably due to error cancelations since the model is too simple for a precise result in general for transition metal complexes. Overall, the FCIQMC-sc-NEVPT2 estimate employing the HF orbitals is slightly large compared to the other results, indicating that orbital optimization plays an important role in constructing a suitable CASCI reference wave function. FCCR also uses the HF orbitals without constructing an active space, being capable of treating the orbital relaxation and higher correlation effects via the exponential ansatz to provide the quantitatively accurate interconversion energies. We cannot pursue further discussions based on the direct comparison due to the difference in the treatment of the basis set and relativistic effects. Nevertheless, the present benchmark results will be useful for assessing high-quality correlated methods for future studies.

To summarize, we have introduced an augmented version of FCCR with a second-order perturbative correction, FCCR(2) that enables a non-iterative treatment of numerous cluster operators with small amplitudes, which are needed for accurate energetics. The computation for $E^{(2)}$ can be efficiently parallelized with a small memory requirement with loops over batches dividing the determinants in the secondary space \mathcal{Q} . The advantage of FCCR over a CI expansion is the use of the exponential ansatz that guarantees an accurate result with a small number of parameters for $\hat{T}_{\mathcal{P}}$ in a size-extensive manner without constructing an active space. This feature leads to a powerful FCCR(2) for obtaining near-FCI solutions by reducing the loads on the second-order correction and subsequent extrapolations. The overall extrapolation distance of FCCR(2)-Ex is only about 1% of the FCI estimate in this work. For open-shell systems, the spin symmetry with unrestricted HF references is systematically recovered by increasing the accuracy of FCCR within the present framework.

Further extensions of the FCCR include calculation of excited states and construction of explicitly correlated wave functions. We leave these research directions for future studies along with discussions on the balance between cost and accuracy against the other methods for strongly correlated systems as partially presented in the case of equilibrium benzene.²⁹

Acknowledgement

This work was supported by MEXT as a “ Priority Issue on Post-K computer (supercomputer Fugaku) ” (Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use) and JSPS Grants-in-Aids for Scientific Research (A) (Grant No. JP18H03900). We are also indebted to the HPCI System Research project (Project ID: hp150228, hp120278, hp160202, hp170259, hp180216, and hp190175), Academic Center for Computing and Media Studies (ACCMS) of Kyoto University, Education Center on Computational Science and Engineering (ECCSE) of Kobe university, and the Research Center for Computational Science, Okazaki, Japan for the use of computer resources.

Supporting Information Available

Details of the working equations and results are collected in Supporting Information.

This material is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- (2) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

- (3) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (4) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference nature of chemistry: The coupled-cluster view. *Chem. Rev.* **2012**, *112*, 182–243.
- (5) Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *WIREs Computational Molecular Science* **2013**, *3*, 176–197.
- (6) Xu, E.; Uejima, M.; Ten-no, S. L. Full coupled-cluster reduction for accurate description of strong electron correlation. *Phys. Rev. Lett.* **2018**, *121*, 113001.
- (7) Liu, W.; Hoffmann, M. R. iCI: Iterative CI toward full CI. *J. Chem. Theory Comput.* **2016**, *12*, 1169–1178.
- (8) Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.
- (9) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112.
- (10) Schriber, J. B.; Evangelista, F. A. Adaptive configuration interaction for computing challenging electronic excited states with tunable accuracy. *J. Chem. Phys.* **2016**, *144*.
- (11) Booth, G. H.; Thom, A. J.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 054106.

- (12) Cleland, D.; Booth, G.; Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **2010**, *132*, 041103.
- (13) Petruzielo, F. R.; Holmes, A. A.; Changlani, H. J.; Nightingale, M. P.; Umrigar, C. J. Semistochastic projector monte carlo method. *Phys. Rev. Lett.* **2012**, *109*, 230201.
- (14) Ten-no, S. Stochastic determination of effective Hamiltonian for the full configuration interaction solution of quasi-degenerate electronic states. *J. Chem. Phys.* **2013**, *138*, 164126.
- (15) Gwaltney, S. R.; Head-Gordon, M. A second-order correction to singles and doubles coupled-cluster methods based on a perturbative expansion of a similarity-transformed Hamiltonian. *Chem. Phys. Lett.* **2000**, *323*, 21–28.
- (16) Hirata, S.; Nooijen, M.; Grabowski, I.; Bartlett, R. J. Perturbative corrections to coupled-cluster and equation-of-motion coupled-cluster energies: A determinantal analysis. *J. Chem. Phys.* **2001**, *114*, 3919–3928.
- (17) GELLAN, A hierarchical massively-parallel quantum chemistry program, Kobe university.
- (18) Tubman, N. M.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. An Efficient Deterministic Perturbation Theory for Selected Configuration Interaction Methods. *arXiv preprint arXiv:1808.02049* **2018**,
- (19) Li, J.; Otten, M.; Holmes, A. A.; Sharma, S.; Umrigar, C. J. Fast Semistochastic Heat-Bath Configuration Interaction. *J. Chem. Phys.* **2018**, *149*, 214110.
- (20) Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern Approaches to Exact Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method. *J. Chem. Theory Comput.* **2020**, *16*, 2139–2159.

- (21) Zhang, N.; Liu, W.; Hoffmann, M. R. Iterative Configuration Interaction with Selection. *J. Chem. Theory Comput.* **2020**, *16*, 2296–2316.
- (22) Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (23) Kucharski, S. A.; Bartlett, R. J. Recursive Intermediate Factorization and Complete Computational Linearization of the Coupled-Cluster Single, Double, Triple, and Quadruple Excitation Equations. *Theor. Chim. Acta* **1991**, *80*, 387–405.
- (24) Klopper, W.; Manby, F. R.; Ten-no, S.; Valeev, E. F. R12 methods in explicitly correlated molecular electronic structure theory. *Int. Rev. Phys. Chem.* **2006**, *25*, 427–468.
- (25) Ten-no, S.; Noga, J. Explicitly correlated electronic structure theory from R12/F12 ansätze. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 114–125.
- (26) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly correlated electrons in molecules. *Chem. Rev.* **2012**, *112*, 4–74.
- (27) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly correlated R12/F12 methods for electronic structure. *Chem. Rev.* **2012**, *112*, 75–107.
- (28) Grüneis, A.; Hirata, S.; Ohnishi, Y.-y.; Ten-no, S. Perspective: Explicitly correlated electronic structure theory for complex systems. *J. Chem. Phys.* **2017**, *146*, 080901.
- (29) Eriksen, J. J.; Anderson, T.; Deustua, J. E.; Ghanem, K.; Hait, D.; Hoffmann, M. R.; Lee, S.; Levine, D. S.; Magoulas, I.; Shen, J.; Tubman, N. M.; Whaley, K. B.; Xu, E.; Yao, Y.; Zhang, N.; Alavi, A.; Chan, G. K.-L.; Head-Gordon, M.; Liu, W.; Piecuch, P.; Sharma, S.; Ten-no, S. L.; Umrigar, C. J.; Gauss, J. The ground state electronic energy of benzene. *J. Phys. Chem. Lett.* **2020**, *11*, 8922–8929.
- (30) Eriksen, J. J.; Gauss, J. Many-body expanded full configuration interaction. I. Weakly correlated regime. *J. Chem. Theory Comput.* **2018**, *14*, 5180–5191.

- (31) Eriksen, J. J.; Gauss, J. Many-body expanded full configuration interaction. II. Strongly correlated regime. *J. Chem. Theory Comput.* **2019**, *15*, 4873–4884.
- (32) Chan, G. K.-L.; Sharma, S. The density matrix renormalization group in quantum chemistry. *Ann. Rev. Phys. Chem.* **2011**, *62*, 465–481.
- (33) Ghanem, K.; Lozovoi, A. Y.; Alavi, A. Unbiasing the initiator approximation in full configuration interaction quantum Monte Carlo. *J. Chem. Phys.* **2019**, *151*, 224108.
- (34) Deustua, J. E.; Magoulas, I.; Shen, J.; Piecuch, P. Communication: Approaching exact quantum chemistry by cluster analysis of full configuration interaction quantum Monte Carlo wave functions. *J. Chem. Phys.* **2018**, *149*, 151101.
- (35) Lee, J.; Malone, F. D.; Reichman, D. R. Note: The performance of phaseless auxiliary-field quantum Monte Carlo on the ground state electronic energy of benzene. *arXiv preprint arXiv:2008.04736* **2020**,
- (36) Loos, P.-F.; Damour, Y.; Scemama, A. Note: The performance of CIPSI on the ground state electronic energy of benzene. *arXiv preprint arXiv:2008.11145* **2020**,
- (37) Cramer, C. J.; Włoch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical models on the Cu₂O₂ torture track: Mechanistic implications for oxytyrosinase and small-molecule analogues. *J. Phys. Chem. A* **2006**, *110*, 1991–2004.
- (38) Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of Cu O₂ and Cu₂ O₂ systems. *J. Chem. Phys.* **2008**, *128*, 204109.
- (39) Phung, Q. M.; Wouters, S.; Pierloot, K. Cumulant approximated second-order perturbation theory based on the density matrix renormalization group for transition metal complexes: A benchmark study. *J. Chem. Theory Comput.* **2016**, *12*, 4352–4361.

- (40) Anderson, R. J.; Shiozaki, T.; Booth, G. H. Efficient and stochastic multireference perturbation theory for large active spaces within a full configuration interaction quantum Monte Carlo framework. *J. Chem. Phys.* **2020**, *152*, 054101.