PCCP

Combined Quantum Monte Carlo – Effective Fragment Molecular Orbital Method: Fragmentation Across Covalent Bonds

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-12-2020-006528.R1
Article Type:	Paper
Date Submitted by the Author:	02-Jun-2021
Complete List of Authors:	Zahariev, Federico; Iowa State University of Science and Technology, Chemistry Gordon, Mark; Iowa State University Department of Chemistry

Combined Quantum Monte Carlo – Effective Fragment Molecular Orbital Method: Fragmentation Across Covalent Bonds

F. Zahariev[a]* and M. S. Gordon[b]*
Department of Chemistry and Ames Laboratory
Iowa State University, Ames, IA 50011

a. federico@si.msg.chem.iastate.edu
b. mark@ si.msg.chem.iastate.edu

The previously developed combined Quantum Monte Carlo – Effective Fragment Molecular Orbital (QMC-EFMO) method is extended to systems in which the fragmentation process cuts across covalent molecular bonds. The extended QMC-EFMO capability is demonstrated on a few model systems, including the glycine tetramer, the diglycine reaction to form a dipeptide, and polyalanine chains of increasing length. The agreement between full QMC and QMC-EFMO for the correlation energy is within 2 kcal/mol and for the correlation energy differences is within 1 kcal/mol.

1. Introduction

The Quantum Monte Carlo (QMC) method is a family of interrelated stochastic approaches, such as the Variational Monte Carlo (VMC) and Diffusion Monte Carlo (DMC) methods, for solving the Schrodinger equation [1]. The statistical error of QMC results can be monitored and gradually lowered by extensions of the simulation time [1]. The current practice of using VMC and DMC in a two-step sequence produces energies with an accuracy that is usually below 1 kcal/mol [2, 3], although deviations in certain cases have been reported [4, 5]. The accuracy is mainly limited by the DMC fixed-node approximation [6, 7].

The QMC approach has had a steady increase in popularity throughout the last two decades, essentially following the increase in the power of supercomputers [1]. Although QMC has a favorable scaling of the computational time with respect to the number of electrons, which is close to cubic scaling [8, 9], the constant of proportionality (pre-factor) in front of the cubic factor is significantly larger than that for the Hatree-Fock (HF) and density functional theory (DFT) methods. Fortunately, it is possible to parallelize the QMC algorithms to massively parallel supercomputers and thus compensate for the relatively large cubic pre-factor [9].

The combination of high accuracy with favorable computational scaling makes QMC an attractive alternative to the more traditional *ab-initio* methods.

The QMC method was recently combined with the effective fragment molecular orbital (EFMO) method [8] in order to further reduce the computational scaling of QMC. Fragmentation methods, frequently based on a many-body expansion, subdivide a large molecular system into a collection of atomic fragments [11]. The fragmentation facilitates the parallelization process, as each fragment can be associated with a separate group of computational nodes, and thus the overall computational cost is reduced to that of the largest fragment. EFMO is a particularly efficient fragmentation method with a reasonable accuracy [12-15].

The QMC-EFMO method is extended to systems in which the fragmentation process cuts across covalent molecular bonds. Both the basic QMC-EFMO method and its novel extension are described in Section 2. The extended QMC-EFMO capability is demonstrated on a few model systems in Section 3 before presenting a summary in section 4.

2. Theoretical background (QMC-EFMO)

QMC-EFMO is a combination of the QMC and EFMO methods. In turn, EFMO is itself a combination of two methods, the fragment molecular orbital (FMO) method [16] and the effective fragment potential (EFP) method [17-19]. In this section, the basic form of the QMC-EFMO method is presented in the first subsection. In addition, the QMC,

FMO, EFP, EFMO, and QMC-EFMO methods are briefly reviewed for completeness of the background presentation. The novel extension of the QMC-EFMO method for fragmentation across covalent bonds is presented in the second subsection.

2.1 QMC-EFMO

In the combined QMC-EFMO method [10], the ground state energy expansion, up to dimer corrections, is given by

$$E^{QMC-EFMO2} = \sum_i E_i^{QMC} + \sum_{i>j}^{R_{i,j} \le R_{cut}} \left(E_{ij}^{QMC} - E_i^{QMC} - E_j^{QMC} - E_{ij}^{pol} \right) + \sum_{i<j}^{R_{i,j} > R_{cut}} \left(E_{ij}^{Coul} + E_{ij}^{disp} + E_{ij}^{ex-rep} + E_{ij}^{ct} \right) + E_{tot}^{pol}, \quad (1)$$

where R_{cut} is the cut-off distance beyond which fragment-fragment interactions are treated by the EFP method; E_i^{QMC} and E_j^{QMC} are the QMC monomer energies, E_{ij}^{QMC} is the QMC dimer energy; E_{ij}^{pol}, E_{ij}^{Coul}, E_{ij}^{disp}, E_{ij}^{ex-rep}, and E_{ij}^{ct} are, respectively, the polarization, Coulomb, dispersion, exchange-repulsion and charge-transfer parts of the EFP dimer energy; and E_{tot}^{pol} is the EFP polarization energy of the entire molecule. The inter-fragment separation $R_{i,j}$ is measured in terms of the van der Waals radii of atoms in fragments I and J [12]. It has been previously demonstrated for water clusters of modest size (e.g., 32 water molecules) that, with virtually no loss in accuracy, R_{cut} can be set at a sufficiently small value, so that nearly all of the dimer interactions are treated with EFP, rather than quantum, methods [13]. Consequently, for large systems with hundreds or even thousands of atoms, only a very small number of dimers have to be computed by quantum (e.g., QMC) methods. In fact, since the quantum dimers are selected by a cut-off criterion around each monomer, their total number is proportional to the total number of monomers.

QMC. In the currently most common QMC procedure, a VMC calculation is followed by a DMC calculation. At the VMC stage, the parameters of a wave function *ansatz* are varied, so that the expectation value of the energy, variance, or a linear combination of the two is minimized [1]. The multidimensional integrals of the expectation values are computed directly by a Monte Carlo approach to integration, rather than through the intermediate use of one- and two-electron integrals as in the traditional quantum-chemistry methods. The VMC wave function *ansatz* usually consists of a single- or multi-determinant wave function, which comes from a preliminary (post-)HF or DFT calculation, multiplied by a Jastrow factor [20]. The expansion coefficients of the molecular orbitals in terms of the computational basis set and the coefficients of the multi-determinant expansion can be used as variational parameters, too. The Jastrow factor contains variational parameters for the electron-nuclear, electron-electron, and possibly electron-electron-nuclear correlation effects [1].

At the next DMC stage, the optimized VMC wave function is used as a guide for the propagation of random walkers that simulate the imaginary-time dynamics of the Schrodinger equation [1]. A basic result of quantum mechanics is that any mixed quantum state, which is a superposition of the ground and excited states, will gradually settle in the ground state, if the mixed state is propagated in imaginary time by the time-dependent Schrodinger equation. The imaginary-time Schrodinger equation can be thought of as a generalization of the usual diffusion equation and thus be simulated by random walkers. Since the fermionic wave function of electrons has both negative and positive lobes, while the density of the random walkers approximate only a positive valued bosonic wave function, the notorious fermion-sign problem has to be addressed [1]. A common solution to the fermion-sign problem in DMC is the fixed-node (FN) approximation [6,7,21-24], which is the main source of QMC errors. In the FN approximation, the nodes of the VMC wave function are used as boundaries between the DMC nodal regions. The DMC random walkers are not allowed to cross the nodal boundaries, which are fixed throughout the entire DMC process. Since DMC is variational and only weakly depends on the basis set as opposed to VMC, the trial wave functions can be screened in order to reduce the error linked to the FN approximation [8, 22].

The VMC-DMC combination produces energies that are on average very accurate, about 1 kcal/mol for chemical systems and a few tens of meV/unit-cell for solid-state systems, an accuracy comparable and sometimes superior to that of the coupled cluster method with singles, doubles, and perturbative triples (CCSD(T)) [3].

In addition, QMC has a favorable scaling of computer time and resources with respect to the system size. If atomic or other spatially localized basis sets are used, the computational scaling of QMC is estimated to be ~(N^3+εN^4), where N is the number of electrons and $\varepsilon \approx 10^{-4}$ is a relatively small number [25, 26]. Although the prefactor of the ~(N^3+εN^4) scaling is nearly three orders of magnitude larger than the scaling prefactors for the Hartree-Fock (HF) and density functional theory (DFT) methods [27, 28], it can be effectively eliminated by an efficient parallelization of the QMC algorithms that scales about linearly with the number of cores [9, 29, 30, 31, 32].

Due to the high accuracy and favorable computational scaling, the QMC method is a promising alternative to the more traditional quantum chemistry computational methods. QMC-EFMO is applicable to excited state calculations as well [10].

FMO. Once the molecular system has been subdivided into fragments, each monomer (an individual fragment) is computed by HF or DFT in the Coulomb field of the other monomers until self-consistency is achieved. Next, dimers (pairs of fragments), trimers (triplets of fragments), and possibly higher-order many-body fragment combinations are computed as corrections to the original self-consistent monomer calculation. The correlation energy can be further computed on the self-consistent monomers, dimers, and trimers by post-HF methods. Up to a trimer correction, the FMO energy is given by

$$E(FMO3) = \sum_i E_i + \sum_{j>i} \left(E_{ij} - E_i - E_j \right) + \sum_{k>j>i} \begin{bmatrix} \left(E_{ijk} - E_i - E_j - E_k \right) - \left(E_{ij} - E_i - E_j \right) \\ - \left(E_{jk} - E_j - E_k \right) - \left(E_{ik} - E_i - E_k \right) \end{bmatrix}, \quad (2)$$

where E_i, E_{ij}, E_{ijk} are the monomer, dimer, and trimer energies, respectively [16].

EFP. Intermolecular interactions can be accurately and efficiently calculated by EFP, which is a semi-classical method directly based on quantum mechanics and has no empirically fitted parameters [17-19]. The EFP energy is derived from many-body perturbation theory and has terms that represent Coulomb effects (E^{coul}), induction/polarization (E^{pol}), exchange repulsion (E^{exrep}), dispersion (E^{disp}), and charge transfer (E^{ct}). EFP is an extensible method that currently has second order Moller-Plesset perturbation theory (MP2) quality for intermolecular interactions, while at the same time being orders of magnitude less computationally expensive [33].

EFMO. The EFMO method is a combination of the two other fragmentation methods described above, FMO and EFP. EFMO improves upon FMO in two significant ways.

First, the monomer energies in EFMO, as opposed to the monomer energies in FMO, are not computed self-consistently. Instead, the total EFP polarization energy is added to the sum of the EFMO monomer energies. The dimer EFP polarization energy is subtracted from the dimer correction to avoid double counting the EFP polarization energy. In fact, the self-consistent EFP polarization energy incorporates many-body interactions so well that the accuracy of the dimer-corrected EFMO can be competitive with that of the expensive trimer-corrected FMO (FMO3).

Second, the highly accurate EFP interaction energy, instead of a simple electrostatic potential energy, is used for interactions between fragments that are separated by more than a cutoff distance.

The improvements of EFMO over FMO in both accuracy and efficiency have been demonstrated in calculations on realistic molecular systems [12].

EFMO is an excellent companion to QMC, as it is an efficient and accurate fragmentation method with a multi-level parallel implementation in the electronic structure chemistry program General Atomic and Molecular Electronic Structure System (GAMESS) [34, 35, 36].

2.2 QMC-EFMO (extension for fragmentation across covalent bonds)

It is useful to observe that the total QMC-EFMO energy by Eq. (1) can be split into two additive components, the Hartree-Fock EFMO energy

$$E^{HF-EFMO2} = \sum_i E_i^{HF} + \sum_{i>j}^{R_{i,j}\le R_{cut}} \left(E_{ij}^{HF} - E_i^{HF} - E_j^{HF} - E_{ij}^{pol} \right) + \sum_{i<j}^{R_{i,j}>R_{cut}} \left(E_{ij}^{Coul} + E_{ij}^{disp} + E_{ij}^{ex-rep} + E_{ij}^{ct} \right) + E_{tot}^{pol} \quad (3)$$

and the QMC-EFMO correlation energy

$$E^{QMC-EFMO2(c.e.)} = \sum_i E_i^{QMC(c.e)} + \sum_{i>j}^{R_{i,j}\le R_{cut}} \left(E_{ij}^{QMC(c.e.)} - E_i^{QMC(c.e)} - E_j^{QMC(c.e)} \right), \quad (4)$$

i.e.

$$E^{QMC-EFMO2} = E^{HF-EFMO2} + E^{QMC-EFMO2(c.e.)}, \quad (5)$$

where the QMC monomer and dimer correlation energies are

$$E_i^{QMC(c.e)} = E_i^{QMC} - E_i^{HF} \quad (6a)$$

and

$$E_{ij}^{QMC(c.e.)} = E_{ij}^{QMC} - E_{ij}^{HF}. \quad (6b)$$

The separation of the terms in Eq. (5) allows the fragmentation of the HF and QMC correlation energies to be treated differently, if needed.

When fragmentation cuts across covalent bonds, FMO and EFMO employ either the Hybrid Orbital Projection (HOP) [37] or Adapted Frozen Orbital (AFO) [38, 39] method at the place of bond breaking. For QMC-EFMO, the HF-EFMO calculation is done

with the AFO method but the QMC-EFMO counterpart uses hydrogen cap atoms at the severed bonds instead of the more sophisticated AFO orbital projection methodology in order to avoid complicated modifications to the Hamiltonian in the QMC calculations. The lengths of the hydrogen-capped bonds are appropriately scaled, while keeping these bond directions fixed. Since the QMC component calculates just the correlation energy, the hydrogen-capped approximation to AFO in QMC is deemed acceptable. The slight loss of precision due to replacing the AFO approach with hydrogen capping is applicable only for the calculation of the correlation energy by QMC. As the illustrative examples in the next section demonstrate, the QMC-EFMO calculations can still successfully recover both the correlation and total energies obtained from full QMC calculations.

GAMESS was used for the HF and EFMO calculations. The QMC calculations for the correlation energies were done using QMCPACK [32]. The integration between GAMESS and QMCAPCK is described in ref. [10].

A double-basis approach was used in the QMC-EFMO calculations, following ref. 40. The EFMO and full Hartree-Fock energies were computed by the 6-311++G(d,p) basis set, while the full and EFMO correlation energies of QMC were computed by the double-zeta Burkatzki-Filippi-Dolg (BFD) effective core potential (ECP) basis set [41]. The optimized geometries of the reactants and transitions state for the diglycine reaction and polyalanine chains were obtained using the B3LYP density functional [42,43,44] and the 6-31+G(d) basis set.

All of the computations were carried out on the Theta computational cluster at Argonne National Laboratory. Each node of Theta has an Intel Knight Landing (KNL) 7230 processor with 16 GB of MCDRAM and 192 GB of DDR4. Each node has 64 cores and each core has 4 SMT hardware threads. The Cray Aries network with 3-level Dragonfly topology is the high-speed interconnect between the nodes of Theta.

3. Illustrative computational examples

Glycine tetramer. The extended QMC-EFMO capability to fragmenting across covalent bonds was first tested on a glycine tetramer molecule that was split into four fragments (see Fig. 1). The correlation energies of glycine tetramer calculated using pure QMC and QMC-EFMO are respectively -2364.90 ± 0.44 kcal/mol and -2366.11 ± 0.10 kcal/mol. The QMC-EFMO correlation energy error is about 1.2 kcal/mol.

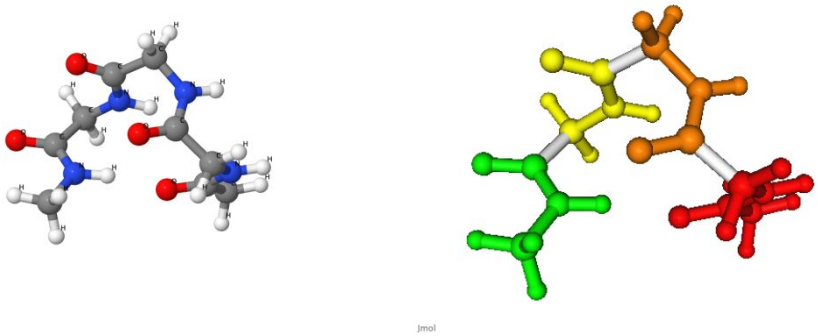

Figure 1. Glycine tetramer. The detailed molecular structure is on the left-hand side and the fragmentation scheme is on the right-hand side, where each color indicates a different fragment.

Dipeptide bond formation. As a further example, consider the reaction of two glycine molecules to form a dipeptide bond [45]. The reaction, Eq. (7), is illustrated in Figure 2.

$$NH_2\text{-}CH_2\text{-}COOH + NH2\text{-}CH_2\text{-}COOH \rightarrow NH_2\text{-}CH_2\text{-}CO\text{-}NH\text{-}COOH \quad (7)$$

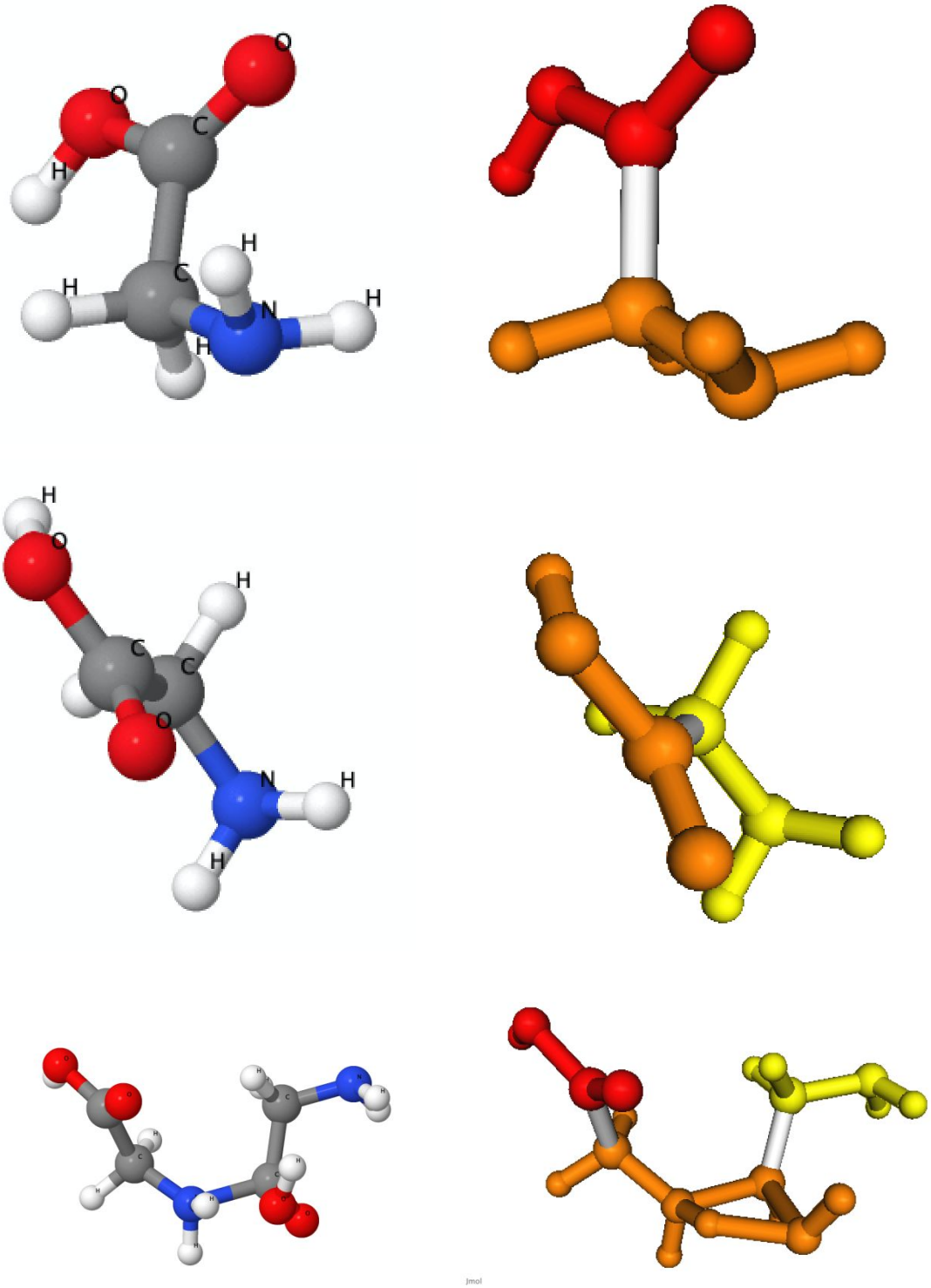

Figure 2. The two glycine molecules (reactants) on the upper part of the figure and the transition state TS3 of the diglycine reaction at the very bottom. The detailed molecular

structures are on the left-hand sides and the fragmentation schemes are on the right-hand sides, where each color indicates a different fragment.

The Hartree Fock, correlation, and total energies (in kcal/mol) of reactants, transition state, and reaction barrier of the diglycine reaction with and without fragmentation by EFMO are shown in Table 1. The QMC-EFMO correlation energy error with respect to the un-fragmented QMC correlation energy is only 0.33 kcal/mol for the reactants, 0.16 kcal/mol for the transition state, and -0.18 kcal/mol for the reaction barrier. The Hartree Fock EFMO energy error with respect to the un-fragmented Hartree Fock energy is -0.02 kcal/mol for the reactants, 1.47 kcal/mol for the transition state, and 1.49 kcal/mol for the reaction barrier. Since the computation time of the Hartree Fock energy is a very small fraction of the QMC-EFMO calculations (See following discussion), one can often afford to compute the Hartree Fock energy in full, without EFMO fragmentation. Thus, Table 1 compares the un-fragmented QMC total energies with the QMC-EFMO total energies, as well as to a modified total QMC-EFMO energy, which is obtained by summing up the total Hartree Fock energy with the QMC-EFMO correlation energy. The QMC-EFMO error is 0.31 kcal/mol for the reactants, 1.63 kcal/mol for the transition state, and 1.31 kcal/mol for the reaction barrier. In contrast, the modified QMC-EFMO error is only 0.33 kcal/mol for the reactants, 0.16 kcal/mol for the transition state, and -0.18 kcal/mol for the reaction barrier.

	reactants	transition state	reaction barrier
EFMO-HF	-355059.41	-354988.72	70.69
HF	-355059.39	-354990.19	69.20
QMC-EFMO corr. en.	-1495.27 ± 0.06	-1520.51 ± 0.05	-25.25 ± 0.08
QMC corr. en.	-1495.60 ± 0.08	-1520.67 ± 0.07	-25.07 ± 0.11
QMC-EFMO tot. en.	-356554.68 ± 0.06	-356509.23 ± 0.05	45.44 ± 0.08
QMC-EFMO corr. en. + HF	-356554.66 ± 0.06	-356510.70 ± 0.05	43.95 ± 0.08
QMC	-356554.99 ± 0.08	-356510.86 ± 0.07	44.13 ± 0.11

Table 1. The Hartree Fock, QMC correlation, and total energies (in kcal/mol) of reactants, transition state, and reaction barrier of the diglycine reaction with and without

fragmentation by EFMO. The first and second rows show the Hartree Fock energy with and without the EFMO fragmentation, respectively. The third and fourth rows show the QCM correlation energy with and without EFMO fragmentation, respectively. The fifth and seventh rows show the QMC total energy with and without EFMO fragmentation, respectively. The sixth row shows the QMC total energy by adding the Hartree Fock energy without fragmentation to the QMC correlation energy with EFMO fragmentation. "Corr. en." stands for correlation energy.

Silica model. For a more complex test of QMC-EFMO fragmentation across covalent bonds, consider a small model of a mesoporous silica nanoparticle (MSN), a highly selective heterogeneous catalyst [46]. For the purpose of the present work, two small silica-based rings (see Figure 3, where the upper and lower parts of the figure correspond to the two silica-based rings) are investigated. The two MSN model systems representing the reactant (R in the upper part of Fig. 3) and product (P in the lower part of Fig. 3) of methane adsorption onto the MSN model:

$$R + CH_4 \rightarrow P + H_2 \qquad (8)$$

Tables 2 and 3 show the energy components of the reactant, product, and the difference between the reactant and product energies (reaction thermodynamics).

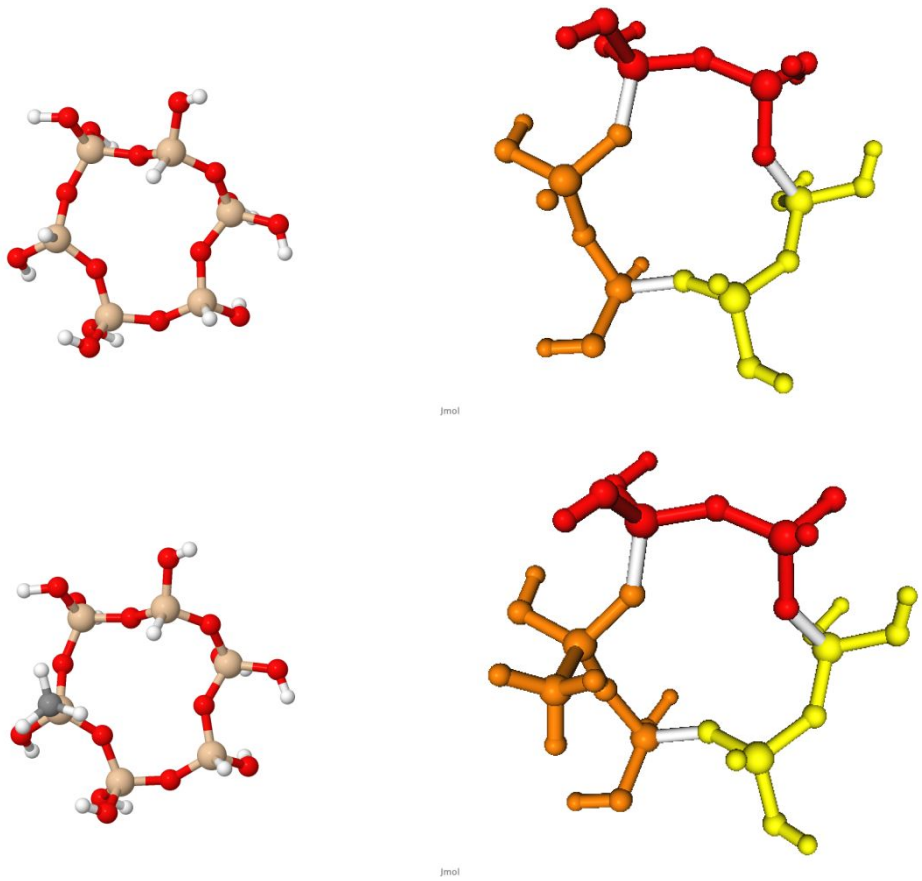

Figure 3. MSN siloxy six-membered ring (reactant R), in the upper part of the figure. The detailed molecular structure is shown on the left-hand side and the fragmentation scheme is shown on the right-hand side, where each color represents a different fragment. MSN siloxy six-membered ring with CH_3 replacing a H in the reactant (product: P). in the lower part of the figure. The detailed molecular structure is on the left-hand side and the fragmentation scheme is on the right-hand side, with each fragment indicated by a different color.

The errors in the QMC-EFMO correlation energies with respect to the QMC full correlation energies (~1.19 kcal/mol for Structure R and ~1.54 kcal/mol for Structure P as seen in Table 2) are smaller than the error in the EFMO Hartree-Fock energies with respect to the full Hartree-Fock energies (-3.08 kcal/mol for Structure R and -3.18 kcal/mol for Structure P).

The EFMO energy difference between the products and reactants (6.93 ± 0.32 kcal/mol) reproduces the full un-fragmented energy difference between the products and reactants (7.18 ± 0.27 kcal/mol) with a deviation of only ~0.25 kcal/mol as seen in Table 3.

In this model, the methane adsorption onto the MSN model is endothermic, probably due in part to the simplicity of the MSN model six-membered siloxy ring and the lack of water solvent molecules. In the experiment [47], methane is adsorbed at any pressure at room temperature, hence the reaction is likely exothermic. Nonetheless, the primary purpose of the current work is to assess the accuracy of the QMC-EFMO method with respect to the full QMC method, rather than to compare the model with the experiment.

	Hartree-Fock	QMC corr. Energy	Total energy
Structure R (EFMO)	-1797765.25	-3275.07 ± 0.23	-1801040.32 ± 0.23
Structure R (FULL)	-1797768.33	-3273.88 ± 0.19	-1801042.21 ± 0.19
Structure P (EFMO)	-1822273.38	-3408.96 ± 0.21	-1825682.34 ± 0.21
Structure P (FULL)	-1822276.56	-3407.42 ± 0.19	-1825683.98 ± 0.19
CH_4	-25232.23	-154.94 ± 0.04	-25387.17 ± 0.04
H_2	-710.65	-27.57 ± 0.04	-738.22 ± 0.04

Table 2. The EFMO and un-fragmented energies of structures R and P, CH_4, and H_2 (in kcal/mol).

	E_{reac}	E_{prod}	E_{prod} - E_{reac}
EFMO	-1826427.49± 0.23	-1826420.56 ± 0.21	6.93 ± 0.32
FULL	-1826429.38 ± 0.19	-1826422.20 ± 0.19	7.18 ± 0.27

Table 3. The EFMO and un-fragmented energies of the reactants (E_{reac}), products (E_{prod}), and their difference (in kcal/mol).

Alanine chains. In order to demonstrate the speedup of QMC-EFMO calculations with respect to the corresponding full QMC calculations, consider polyalanine chains of length 4 (PA4), 6 (PA6), and 8 (PA8) (see Fig. 4). PA4, PA6, and PA8 are subdivided into 4, 6, and 8 fragments, respectively, by cutting across the C-N bonds between the alanines. The "speedup" is defined by dividing the QMC time by the QMC-EFMO time.

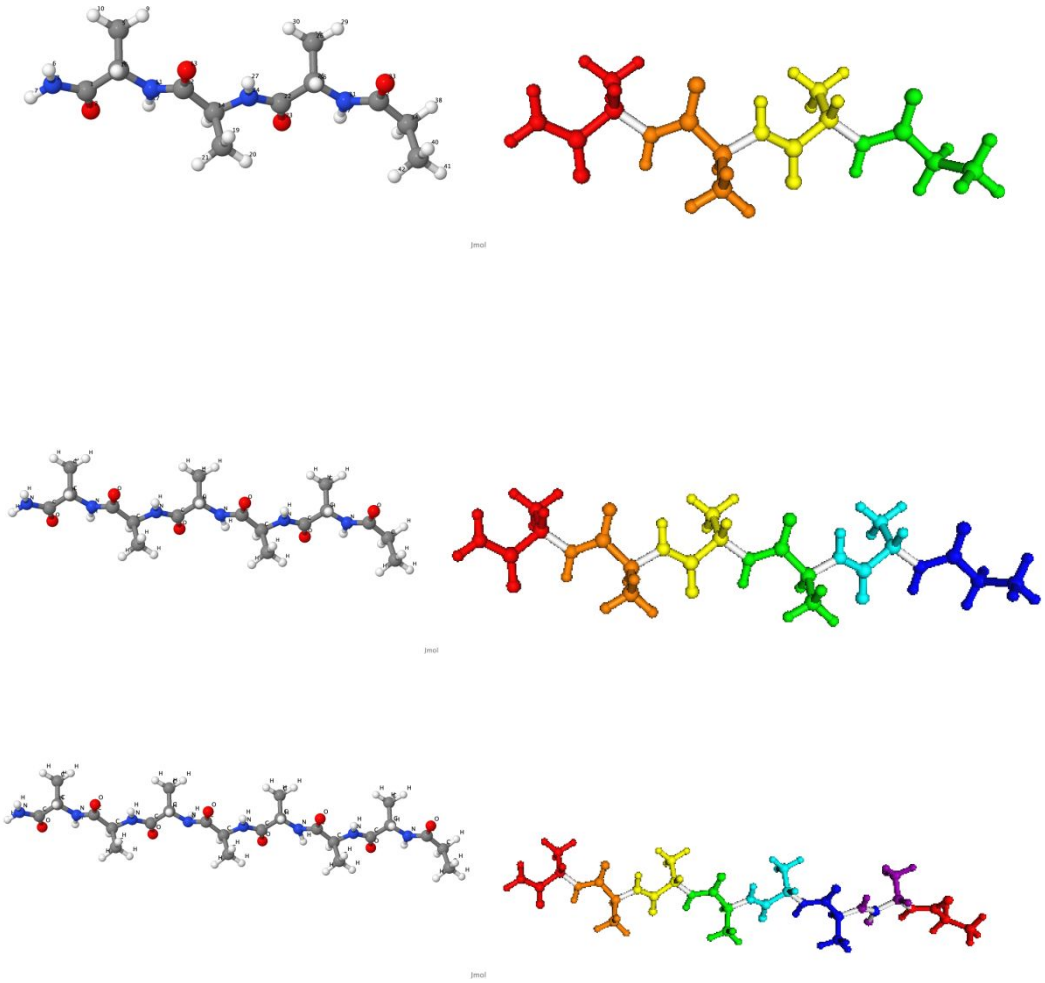

Figure 4. Polyalanine chains of length 4 (PA4), 6 (PA6), and 8 (PA8), from top to bottom. The detailed molecular structures are on the left-hand sides and the fragmentation schemes are on the right-hand sides, where each color indicates a different fragment.

The correlation energies (in kcal/mol) and timings (in node hours) and speedups of QMC-EFMO and un-fragmented QMC for PA4, PA6, and PA8 polyalanine chains are shown in Table 4.

The speedup of QMC-EFMO with respect to QMC steadily increases as the length of the polyalanine chains lengthen and consequently, the number of fragments grows. It is expected that the trend will continue as the number of fragments grows further. The time for the Hartree Fock calculations is a small fraction of the overall computation time. For example, for PA8, the un-fragmented Hartree Fock calculation requires about 8 node hours and the EFMO Hartree Fock calculation requires about 2 node hours, while the total QMC and QMC-EFMO calculations are respectively 10752.0 and 2764.5 node hours.

	PA4	PA6	PA8
corr. en. QMC-EFMO	-2774.41 ± 0.19	-4151.49 ± 0.38	-5527.82 ± 0.21
corr. en. QMC	-2774.27 ± 0.09	-4151.81 ± 0.15	-5528.50 ± 0.09
corr. en. error	-0.14 ± 0.21	0.32 ± 0.46	0.68 ± 0.23
node hours QMC-EFMO	569.6	1202.1	2764.5
node hours QMC	1228.8	3010.6	10752.0
speedup	2.16	2.50	3.89

Table 4. The correlation energies (in kcal/mol) and timings (in node hours and speedups) of QMC-EFMO and un-fragmented QMC for PA4, PA6, and PA8 polyalanine chains.

4. Conclusions

The fragmentation-based QMC-EFMO [10] method has previously been shown to reproduce quite well the full QMC ground and excited state energies, while having an almost linear computational cost. In this work, QMC-EFMO is extended to fragmentation cuts across covalent bonds. The QMC-EFMO application to covalent bond fragmentation was illustrated on the correlation energy of the glycine tetramer molecule and the energy difference obtained for the adsorption of a methane molecule onto a six-membered siloxy ring chosen as a model to represent a mesoporous silica nanoparticle. The comparisons between the full QMC and QMC-EFMO methods for the correlation and total energies demonstrate the good quality for the application of the QMC-EFMO method to fragmentation across covalent bonds.

The demonstration of a successful QMC-EFMO extension to fragmentation across covalent bonds paves the way to larger-scale applications with any type of fragmentation that are expected to also have a close to linear computational scaling with respect to the system size, as in the original QMC-EFMO approach [10].

Acknowledgement. This work was supported by a Department of Energy Exascale Computing Project under Project Number 17-SC-20-SC, administered by the Ames Laboratory.

References

1. B. M. Austin, D. Yu. Zubarev, and W. A. Lester Jr, *Chem. Rev.* 2012, **112**, 263.

2. M. A. Morales, J. McMinis, B. C. Clark, J. Kim, and G. E. Scuseria, *J. Chem. Theory Comput.*, 2012, **8**, 2181.

3. W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys*., 2001, **73**, 33.

4. J. Lee, F. D. Malone, D. R. Reichman, J. Chem. Phys., 2020, **153**, 126101.

5. J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar, J. Gauss, J. Phys. Chem. Lett., 2020, **11**, 8922.

6. J. B. Anderson, *J. Chem. Phys.*, 1976, **65**, 4121.

7. P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester Jr., *J. Chem. Phys.*, 1982, **77**, 5593.

8. M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, J. Chem. Phys., 2016, **144**, 151103.

9. K. P. Esler, J. Kim, D. M. Ceperley, W. Purwanto, E. J. Walter, H. Krakauer, S. W. Zhang, P. R. C. Kent, R. Hennig, C. Umrigar, M. Bajdich, J. Kolorenc, L. Mitas, and A. Srinivasan, *J. Phys. Conf. Ser.*, 2008, **125,** 012057: 1.

10. F. Zahariev and M. S. Gordon, *Mol. Phys.*, 2019, **117**, 1532.

11. M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, *Acc. Chem. Res.*, 2012, **112**, 632.

12. C. Steinmann, D. G. Fedorov, and J. H. Jensen, *J. Phys. Chem. A*, 2010, **114**, 8705.

13. S. R. Pruitt, C. Steinmann, J. H. Jensen, and M. S. Gordon, *J. Chem. Theory Comp.*, 2013, **9**, 2235.

14. S. R. Pruitt, C. Bertoni, K. R. Brorsen, and M. S. Gordon, *Acc. Chem. Res.*, 2014, **47**, 2786.

15. C. Bertoni and M.S. Gordon, *J. Chem. Theory Comp.*, 2016, **12**, 4743.

16. K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, *Chem. Phys. Lett.* **313**, 701 (1999).

17. J. H. Jensen and M. Gordon, *Mol. Phys.*, 1996, **89**, 1313.

18. M. S. Gordon, L. Slipchenko, H. Li, and J. H. Jensen, *Annu. Rep. Comput. Chem.*, 2007, **3**, 177.

19. M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, *J. Phys. Chem. A*, 2001, **105**, 293.

20. R. Jastrow, *Phys. Rev.*, 1955, **98**, 1479.

21. W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, *Phys. Rev. B*, 1999, **60**, 4558.

22. M. Dubecky, *J. Chem. Theory Comput.*, 2017, **13**, 3626.

23. P. G. Hipes, *Phys. Rev. B*, 2011, **83**, 195118.

24. F. Zahariev. M. S. Gordon, and M. Levy, *Phys. Rev. A*, 2018, **98**, 012144.

25. M. D. Towler, *Phys. Stat. Sol. (b)*, 2006, **243**, 2573.

26. W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.*, 2001, **73**, 33.

27. M. Towler, *Quantum Monte Carlo: Accuracy, generality, and scalability in the electron correlation problem*, <http:/vallico.net.casinoqmc>.

28. Frank Jensen, *Introduction to Computational Chemistry* (John Wiley & Son Ltd., 1998).

29. K. P. Esler, J. Kim, D. M. Ceperley, W. Purwanto, E. J. Walter, H. Krakauer, SW Zhang, P. R. C. Kent, R. Hennig, C. Umrigar, M. Bajdich, J. Kolorenc, L. Mitas, and A. Srinivasan, *Journal of Physics: Conference Series*, 2008, **125**, 012057: 1.

30. A. Scemama, M. Caffarel, E. Oseret, and W. Jalby, *J. Comp. Chem.*, 2013, **34**, 938.

31. M. D. Towler, *Psi-k Newsletter*, 2003, **60**, 166.

32. J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali, M. C. Bennett, M. A. Berrill, N. S. Blunt, E. J. L. Borda, M. Casula, D. M. Ceperley, S. Chiesa, B. K. Clark, R. C. Clay III, K. T. Delaney, M. Dewing, K. P. Esler, H. Hao, O. Heinonen, P. R. C. Kent, J. T. Krogel, I. Kylänpää, Y. W. Li, M. G. Lopez, Y. Luo, F. D. Malone, R. M. Martin, A. Mathuriya, J. McMinis, C. A. Melton, L. Mitas, M. A. Morales, E. Neuscamman, W. D. Parker, S. D. P. Flores, N. A. Romero, B. M. Rubenstein, J. A. R. Shea, H. Shin, L. Shulenburger, A. F. Tillack, J. P Townsend, N. M. Tubman, B. Van Der Goetz, J. E. Vincent, D. C. M. Yang, Y. Yang, S. Zhang, and L. Zhao, *Condens. Matter*, 2018, **30**, 195901.

33. J. C. Flick, D. Kosenkov, E. G. Hohenstein, C. D. Sherrill, and L. V. Slipchenko, *J. Chem. Theory Comput.*, 2012, **8**, 2835.

34. 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. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347.

35. M. S. Gordon and M. W. Schmidt, *Advances in electronic structure theory: GAMESS a decade later*, in *Theory and Applications of Computational Chemistry: the first forty years*, C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (editors), Elsevier, Amsterdam, 2005), pp. 1167-1189.

36. G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. G. Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev, M. S, Gordon, *J. Chem. Phys.*, 2020, **152**, 154102.

37. T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi M, and K. Kitaura, *Chem. Phys. Lett.*, 2000, **318**, 614.

38. D. Fedorov, J. Jensen, R. Deka, K. Kitaura, *J. Phys. Chem. A*, 2008, **112**, 11808.

39. D. Fedorov, P. Avramov, J. Jensen, K. Kitaura, *Chem. Phys. Lett.*, 2009, **477**, 169.

40. D. G. Fedorov and K. Kitaura, *Chem. Phys. Lett.*, 2014, **597**, 99.

41. M. Burkatzki, C. Filippi, and M. Dolg, *J. Chem. Phys.*, 2007, **126**, 234105.

42. A. D. Becke, Phys. Rev. A, 1988, **38**, 3098,

43. C. Lee, W. Yang, R. G. Parr, Phys. Rev. B, 1988, **37**, 785; S. H. Vosko, L. Wilk, M. Nusair, Can. J. Phys., 1980, **58**, 1200.

44. A. D. Becke, J. Chem. Phys., 1993, **98**, 5648.

45. J. H. Jensen, K. K. Baldridge, M. S. Gordon, J. Phys. Chem., 1992, **96**, 8340.

46. I. I. Slowing, J. L. Vivero-Escoto, B. G. Trewyn, V. S.-Y. Lin, *J. Matter. Chem.*, 2010, **20**, 7924.

47. C.-C. Liu, H.-J. Chou, C.-Y. Lin, D. Janmanchi, P.-W. Chung, C.-Y. Mou, S. S.-F. Yu, S. I. Chan, *Micropor. Mesopor. Mat.*, 2020, **293**, 109793.