Summer 2017

Application of quantum Monte Carlo methods to molecular potential energy surfaces

Andrew Douglas Powell

Department: Chemistry

APPLICATION OF QUANTUM MONTE CARLO METHODS TO MOLECULAR POTENTIAL ENERGY SURFACES

By

ANDREW DOUGLAS POWELL

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSPHY

in

CHEMISTRY

2017

Approved
Richard Dawes, Advisor
Jeffrey Winiarz
Garry Grubbs
Klaus Woelk
Paul Parris

PUBLICATION DISSERTATION OPTION

This dissertation has been prepared using the publication option. It consists of the following three articles which have been submitted for publication, or will be submitted for publication and have been formatted in the style used by each journal:

Paper I: Pages 15-39 have been accepted into the Journal of Physical Chemistry A.

Paper II: Pages 40-63 have been accepted into the Journal of Chemical Physics.

Paper III: Pages 64-88 are intended for submission into the Journal of Chemical Physics.

ABSTRACT

Various computational methods have been used to generate potential energy surfaces, which can help us simulate and interpret how atoms or molecules behave during a chemical reaction. For accurate work, *ab initio* wavefunction methods have traditionally been used, which have some disadvantages. For example, highly accurate methods scale poorly with system size (n^7 or higher) and are mostly not well parallelized for calculations with multiple processors. One alternative method that has more favorable scaling with system size and is well parallelized is a computational technique called quantum Monte Carlo (QMC). QMC methods scale with the number of electrons as n^3 and have been found to scale almost linearly with the number of processors, even beyond 500,000 cores. However, despite the favorable scaling towards large systems, the cost of QMC methods is relatively expensive for small systems. Small systems nevertheless make important benchmarks necessary for the new methods to gain acceptance. Thus, it was determined to study QMC methods in a few benchmark systems in order to assess its accuracy and routine applicability.

It was found that QMC methods can be very accurate comparing well with experimental measurements and other high-level *ab initio* methods. Benchmark calculations with QMC produced realistic spectroscopic parameters for CO and N_2. However, for small system sizes, they are relatively very expensive to perform with the cost being orders of magnitude higher than traditional methods. Consequently, their use in small systems will likely most often be restricted to only a few geometrical points of interest, unlike traditional methods. Nevertheless, deep insight into the electronic structure of a system can be obtained.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Richard Dawes for mentoring me throughout my program. His encouragement and support have constantly pushed me to give maximum effort. He has shown me the amount of effort, time, and work necessary to pursue my dreams and desires. I would also like to thank the Department of Energy and the National Science Foundation for providing the funding for my research.

I would also like to thank each member of my committee, Dr. Jeffrey Winiarz, Dr. Gary Grubbs, Dr. Paul Parris, and Dr. Klaus Woelk for taking the time to participate on my committee and investing in my future. I would like to thank Dr. Winiarz for helping me in adjusting to my graduate level studies, I would like to thank Dr. Grubbs for encouraging me in my work, Dr. Woelk for ensuring that I know my background, and I Dr. Parris for instructing me in quantum mechanics. I would especially like to thank Drs. Vladimir Tyuterev and Fabien Gatti for training and accommodations in France. I would also like to thank the staff in the chemistry department, Tina, Mel, Shannon, and Alisa, for providing encouragement and support. I would like to thank my fellow graduate students and post-docs for interesting discussions various projects and topics.

Lastly, I would also like to thank family: my parents Douglas and Kimberly, my sisters Vanessa and Veronica, and my brother Aaron. Their love, support, encouragement, thoughts and prayers have been a solid rock for me in my graduate studies. Many nights they have called just to talk and let me know how much they loved me and were thinking about me.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

SECTION

PAPER I

PAPER II

PAPER III

LIST OF TABLES

1. INTRODUCTION

Chemistry is the branch of science investigating the properties, transformation, kinetics, dynamics, *etc*. of molecules. Molecules are made up of atoms, which in turn are composed of charged particles, positively charged nuclei and negatively charged electrons. Theoretical chemistry is the branch of chemistry that applies the laws of physics and mathematics to the study of molecular systems. Computational chemistry implements these theories into algorithms to be solved using computational resources. Various properties and phenomena of systems can be studied, a few examples being relativistic energies, photodissociation, vibrational frequencies, NMR coupling constants, dipole moments, band gaps, surface reactions, and pharmaceutical drug design. Quantum chemistry applies quantum mechanics to study and investigate molecular and atomic systems. For many applications, the fundamental equation governing these interactions is the time-independent Schrödinger equation (SE),

$$H\Psi = E\Psi \tag{1}$$

in which H is the Hamiltonian operator, Ψ is the wavefunction describing the positions of all the fundamental particles for the system of interest, and E is the eigenvalue of the solution to the Hamiltonian. The Hamiltonian operator H is the sum of the potential energy operator V and kinetic energy operator T, *i.e.*,

$$H = T + V \tag{2}$$

The Coulomb Hamiltonian can be written explicitly as the interactions of charged particles, *i.e.*, the nuclei and the electrons, which includes electron-electron repulsion, nuclear-electron attraction, and electron-electron repulsion terms,

$$H = \frac{-1}{2}\sum_i \nabla_i^2 - \sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{A>B}\frac{Z_A Z_B}{R_{AB}} + \sum_{i>j}\frac{1}{r_{ij}} \quad (3)$$

where $\nabla_i^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. If the Schrodinger equation can be solved exactly, then the total energy and the exact wave function of the system will be known.

Due to the many terms and structure of the Coulomb Hamiltonian, it is very difficult to solve due to the motion of the nuclei and electrons being coupled. To simplify the situation, the Born-Oppenheimer (BO) approximation is often introduced, which decouples the motion of the nuclei and the electrons. It is often rationalized by the large relative mass ratio of the nuclei to the electrons with the mass of an electron being 9.109 x 10^{-31} kg and the mass of a proton being 1.727 x 10^{-27} kg. Since the electrons adjust more quickly than the nuclei, their motions can be separated. In the BO approximation, the nuclei are fixed at a series of geometric configurations, and the electronic energy is solved, which in turn provides the potential for the nuclear dynamics.

1.1 ELECTRONIC STRUCTURE THEORY

Brief descriptions of a few pertinent methods in the hierarchy of electronic structure theory are presented. This is by no means intended to be a comprehensive or exhaustive description of the various levels of theories. The interested reader is encouraged to examine the references for a more detailed account.

1.1.1 Hartree-Fock (HF) Method. The simplest wave function model in *ab initio* electronic structure theory is the Hartree-Fock (HF) method. This method primarily serves as the reference configuration for more accurate and sophisticated treatments in electronic structure.

In the HF method, the electronic wave function is approximated by a single configuration (determinant) of spin orbitals, and the energy and the orbitals are optimized with respect to variations of these spin orbitals. The HF ground state is obtained by an iterative procedure where the energy and orbitals are minimized with respect to the orbital rotations. By solving a set of one-electron Schrödinger equations for the spin orbitals, an optimal determinant may be found, which is called the Hartree-Fock determinant. The associated Hamiltonian to solve these equations is called the Fock operator:

$$f = \sum_{PQ} f_{PQ} a_P^\dagger a_Q \tag{5}$$

where the elements f_{PQ} compose the Fock matrix. The Fock operator replaces the two-electron interaction by an effective one-electron Fock potential V:

$$f = h + V \tag{6}$$

which treats the Coulomb repulsion of the electrons in an averaged way. The equations are solved by diagonalizing the Fock matrix. The resulting eigenvectors and eigenvalues are called the canonical spin orbitals and the orbital energies, respectively, of the system. An iterative procedure is used to solve the Fock matrix, and this procedure is called the self-consistent field (SCF) method and, the resulting wave function, the SCF wave function.

To describe the orbitals, basis sets are used, which provides a flexible set of functions. In the limit of an infinite basis set, the complete basis set (CBS) limit is reached. As is shown in Figure 1, there is a hierarchy of different methods that show increasing accuracy. Full configuration interaction with complete basis set demands a lot

of computational resources with computational time, memory, and data storage, making the method infeasible for routine use. The memory, storage, and I/O also increases with increasing basis size. Therefore, a decision must be made to the level of theory and also to the basis set in order to provide sufficiently accurate results for the chemical system of interest.

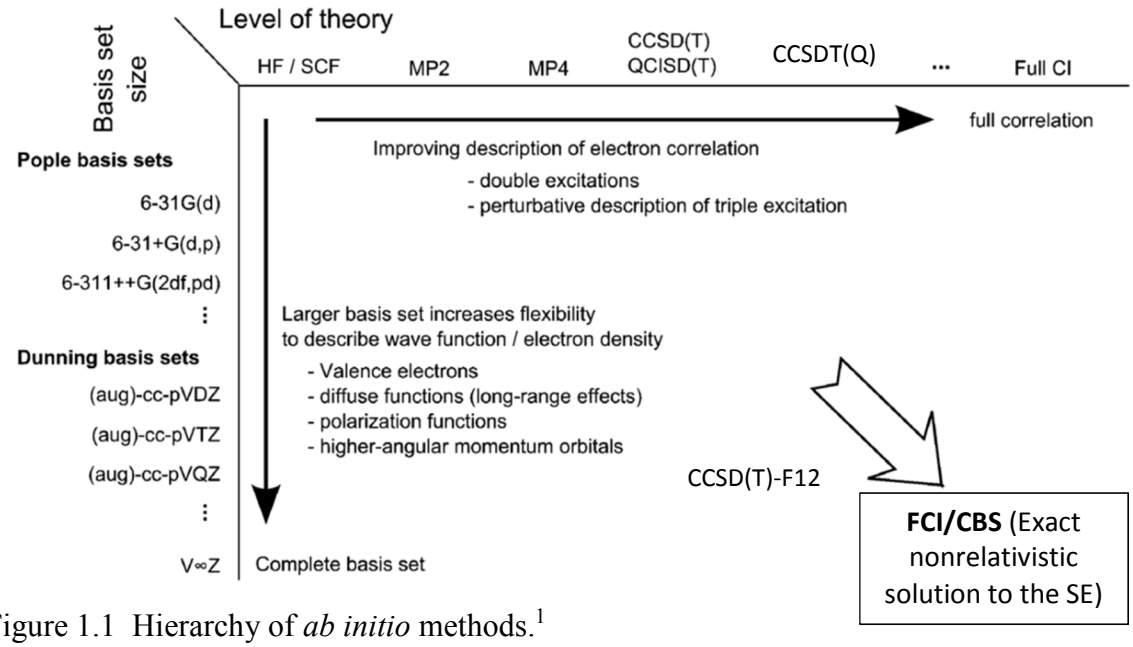

Figure 1.1 Hierarchy of *ab initio* methods.[1]

At an early stage,[2] it was known that, while the HF method can capture a large amount of the total energy of a system (~99%), it is limited in the accuracy of its description of molecular systems. Due to the approximation of average Coulomb electronic repulsion, the instantaneous interactions (correlations) of the electrons are not treated, which are crucial for the description of chemical bond formation.[3] The electron correlation captured by FCI is often separated into the strong (or static correlation) and the dynamic (or weak) correlation. A shortcoming of the HF method can be seen in the example of the H_2 molecule. The HF method incorrectly calculates the dissociation

energy by ~190 kcal/mol (due to neglect of both types of correlation). At the equilibrium bond distance, the Hartree-Fock wave function reasonably describes the most relevant configuration from the exact FCI wave function.

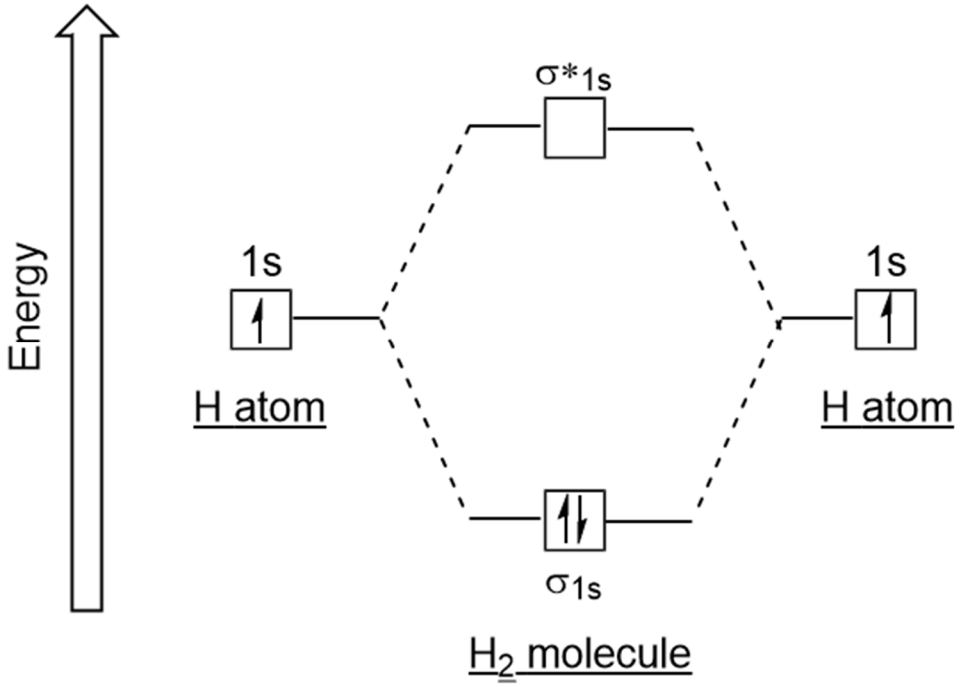

Figure 1.2 Molecular orbital diagram for H_2.[4]

However as the H-H bond distance increases, the HF wave function becomes physically unreasonable, since, at longer distances, each electron should but can't localize on one H atom. The reason is that, as the H-H bond distance increases, the single Hartree-Fock determinant is incapable of accurately treating the evolution in the exact wave function due to bond breaking. A point needs to be made about the different types of electron correlation. For the stable H_2 molecule at equilibrium geometry, the need arises to describe dynamic correlation, the correlated motion of electrons, which comes from the instantaneous Coulomb repulsion of the electrons. In the limit of molecular dissociation, the configurations arising from the degeneracy, or near degeneracy, of the

bonding and antibonding configurations must be treated. This static (or strong) correlation arises because these configurations interact strongly, cannot be treated in isolation, and are unrelated to the repulsion between the electrons.

In order to capture the electron correlation energy, more sophisticated methods are employed. These are called post-Hartree-Fock methods, because the HF method is used as the starting point for the initial orbitals and Slater determinant. The correlation energy is defined as the difference between the energy obtained in the exact solution of the nonrelativistic solution Schrödinger equation, *i.e.*, the full configuration interaction (FCI) energy, and the Hartree-Fock energy and can be written mathematically as

$$E_{correlation} = E_{exact} - E_{HF} \tag{7}$$

It is very accurate, but can only be solved for systems with few (<12 electrons) and limited to small basis sets. The computer time required to reach the FCI solution for larger systems is prohibitive. Therefore, in solving for the electron correlation energy, there is a balance accuracy and computational cost (CPU time). A few of these methods are highlighted in the following sections.

1.1.2 Multiconfigurational Self-Consistent Field (MCSCF). One method to improve HF is to allow for more configurations within the electronic wavefunction in what is called the multi-configurational self-consistent field (MCSCF) method. Similar to HF, MCSCF uses the SCF method to converge to a set of molecular orbitals. However, instead of producing a single Slater determinant, MCSCF gives a reference wave function which includes all possible configurations from a group of active orbitals (see Figure 2), corresponding to perhaps many configurations and thus allowing for the

treatment of static correlation. The wave function is written as a linear combination of these determinants, and, applying the variational method and minimizing the energy, the weights of these configurations are optimized simultaneously with the molecular orbitals.

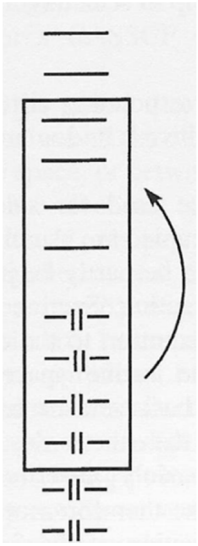

Figure 1.3 MCSCF illustration of excitations within a set of valence orbitals.[5]

MCSCF gives more flexibility to the wave function in describing bonded systems. The reference space is chosen so that the underlying wave function can be qualitatively correct with respect to important valence correlation contributions such as excited states and bond breaking due to dissociation. During dissociation of a diatomic molecular system, the configurations describing the wave function are continuously changing from the equilibrium geometry all the way to separate atomic configurations, and MCSCF allows the important configurations to be determined. Whereas HF gave an unreasonable description as the bond of H_2 was stretched, MCSCF improves description of the wavefunction and gives at least qualitatively correct behavior toward dissociation, *e.g.*, see Figure 1.3. MCSCF allows for the treatment of static correlation, the correlation

arising from degenerate or near-degenerate electron configurations.[6] However, it does not fully account for the instantaneous interactions of electrons (dynamic correlation). To obtain higher accuracy, MCSCF is applied as a reference method for further excitation

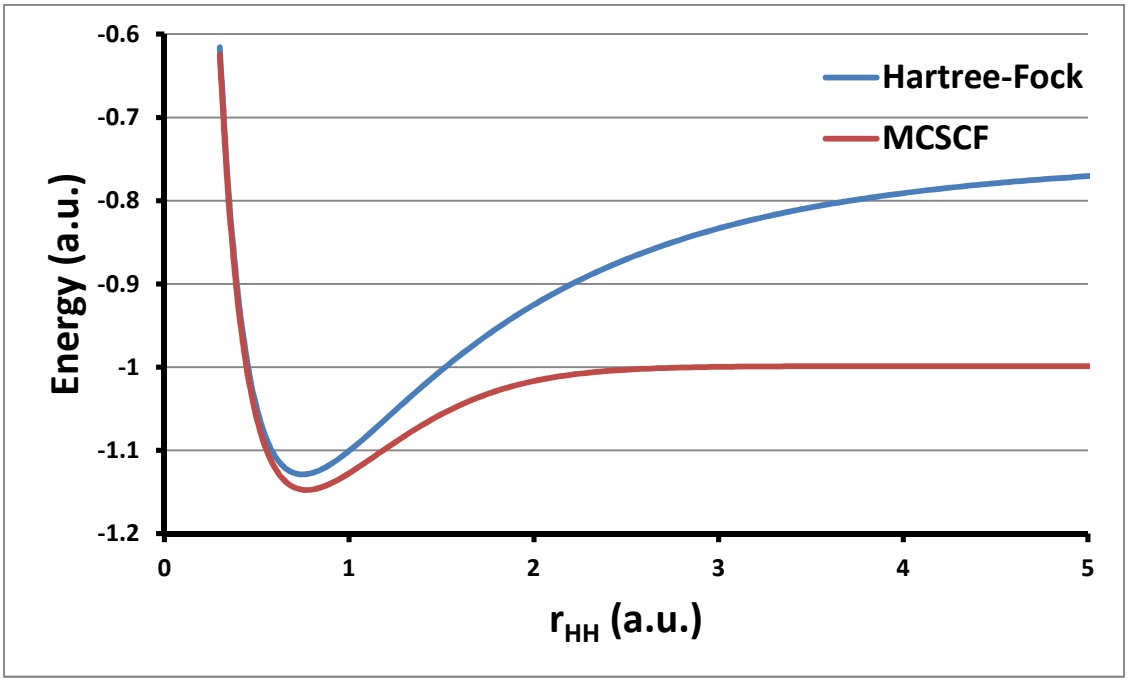

Figure 1.4 Qualitative Comparison of the Hartree-Fock and MCSCF methods for H_2.[5]

treatments which describe dynamic correlation, some of which are described in the following sections. (Strictly speaking some dynamic correlation is obtained in MCSCF, so the classification of dynamic correlation is less rigorous for MCSCF than eq. 4 which applies to single determinant descriptions).

1.1.3 Coupled-Cluster (CC). For systems with a single, dominant electronic configuration, dynamic electron correlation can be well-described by using the coupled-cluster (CC) method, which uses HF as the reference.[7] Mathematically, the CC method begins with an exponential form of the wave function

$$\Psi = \exp(T)\Psi_0 \qquad (5)$$

where $T = T_1 + T_2 + \cdots$. A basic coupled-cluster method is CCSD, coupled-cluster with single and double excitations only, *i.e.*, $T = T_1 + T_2$. With full excitations, the FCI energy is recovered, but since the computational expense is too large for practical use, the excitation level of couple cluster must be truncated. CCSD is often not sufficient for highly accurate work, and the next level of excitation, coupled-cluster with full single, double, and triple excitations (CCSDT), is, while more accurate, is much more computationally demanding. For instance, with system size in which n is the number of electrons, CCSD scales as n^6 whereas CCSDT scales as n^8. To have a method that is more accurate than CCSD but less computationally expensive than CCSDT, a hybrid method was developed: coupled-cluster with single, double, and perturbative contributions of connected triple excitations [CCSD(T)], which scales as n^7. This method is considered the "gold standard" of quantum chemistry due to being more practical than CCSDT and yet still being a highly accurate approximation for computational study.[8]

1.1.4 Multireference Configuration Interaction (MRCI). Multireference configuration interaction (MRCI)[9,10] allows for the calculation of additional dynamic correlation for those systems with multiple important configurations. The reference space for the MRCI method is usually taken from MCSCF. The orbitals and determinants from the MCSCF method are used to construct the initial wave function. A common method is MRCI with single and double excitations of electrons from the active space called MRCI-SD. More excitations can be used but these become extremely expensive to evaluate and very rapidly increases the memory storage and usage of a single or multiple processors. To include some contributions of higher excitations (and approximately restore size consistency), a Davidson correction[11] can be used, which is

abbreviated as (Q). There are several types of available theories but they are beyond the scope of this work.

1.1.5 Moore's Law. In 1965, Gordon Moore observed that the number of components in an integrated circuit had doubled since 1959, and he predicted that this process would continue in the future. In 1975, he later amended this prediction by stating that the semiconductor components would double every two years.[12]

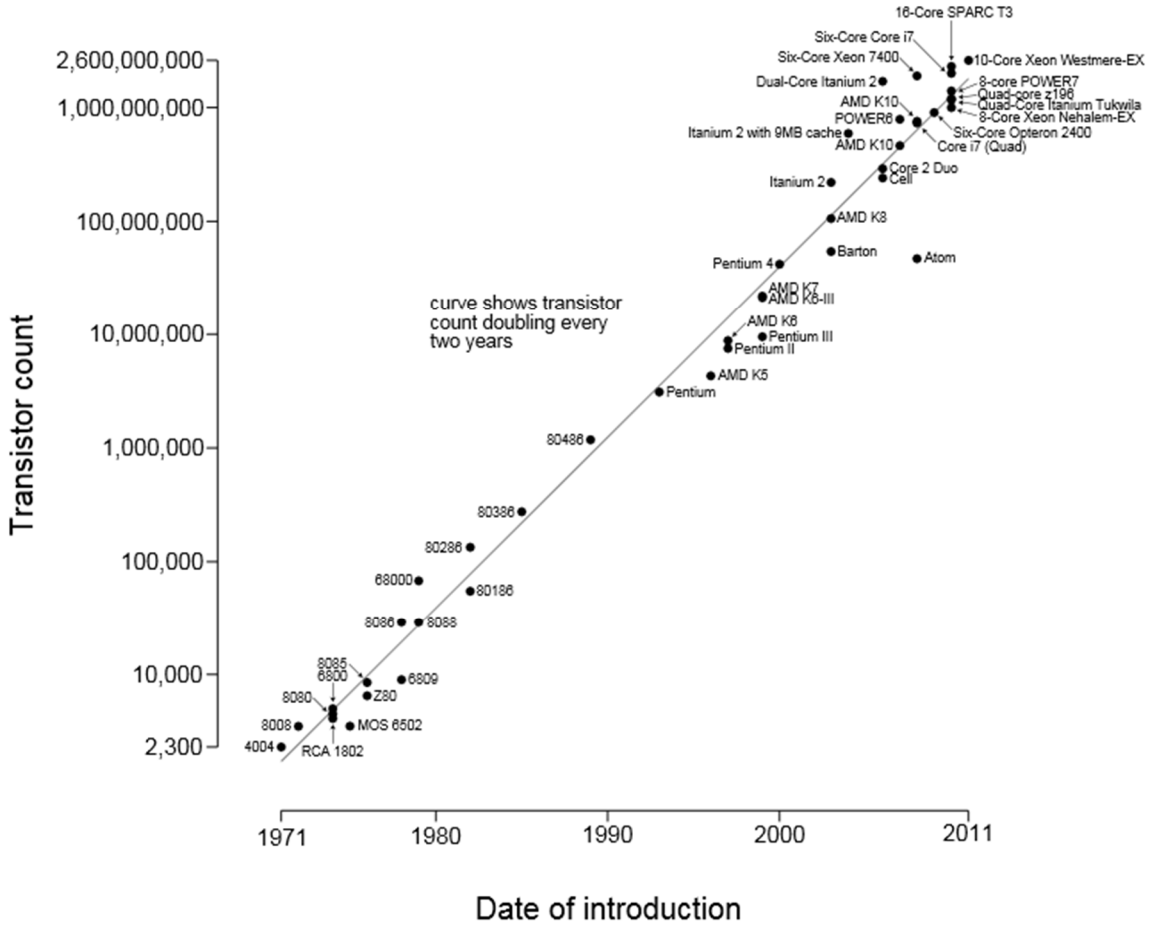

Figure 1.5 A plot of CPU transistor counts vs. dates of introduction.[13]

As the number of components has increased, the speed of these microchips has increased as well. Faster microchips enable the processors to work much more quickly,

rapidly expanding the role of computational chemistry. Theoretical chemists have developed theories and algorithms to give great insight into the electronic structure problem. More recently, microchips are not increasing as much in speed, and a popular idea has been to increase the number of processors used for computation (parallel processing). Many ab initio methods are not well parallelized and are thus unable to take advantage of state-of-the-art computational resources. Also, as has been discussed, the scaling with the number of electrons is poor for many methods. One method for which there is better scaling with system size and well parallelized algorithms is quantum Monte Carlo.

1.2 MONTE CARLO METHODS

A brief description and example is given of the Monte Carlo method. This is then followed by highlighting two of the Quantum Monte Carlo methods used in this work.

1.2.1 Monte Carlo (MC). Monte Carlo methods are algorithms that involve random sampling to obtain numerical results.[14,15] One such example is to evaluate pi using the Monte Carlo method.

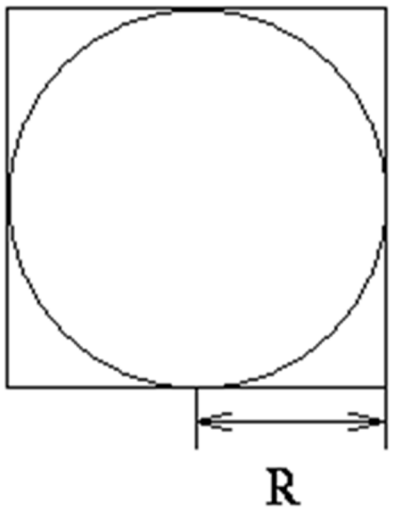

Figure 1.6 Circle of radius R inscribed in a square.

To determine the value of pi, points are first randomly placed in the square. Then ratio of the points found inside and outside the circle can be used to estimate pi . For example, $Area_{circle} = \pi R^2$ and the $Area_{square} = 4R^2$. If these are rearranged to solve for pi, then pi can be defined as

$$\pi = 4\frac{Area_{circle}}{Area_{square}} \tag{6}$$

This means that the ratio of the points sampled inside the circle to those inside the square. With many points sampled, then an accurate value of pi will be approached.

The Monte Carlo method can also be used to estimate the value of an integral. Consider an integral I to be evaluated (equation 7):

$$I = \int f(x)dx \tag{7}$$

This integral can be approximated with the Monte Carlo method by:

$$I \approx D_n = V\frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{8}$$

In the limit of infinite sampling, the value of the integral is obtained (equation 9):

$$\lim_{N\to\infty} D_N = I \tag{9}$$

Of course, infinite sampling is impractical, but, if a finite number of points are used, the integral can be estimated by statistical averaging. This problem alone isn't very useful, but the method can be applied to a wide range of problems and becomes very important for calculating the value for multi-dimensional integrals.

1.2.2 Quantum Monte Carlo (QMC). Applying the Monte Carlo methods to quantum chemistry is known as Quantum Monte Carlo.[16,17] These methods scale with the number of electrons as n^3 whereas traditional high accuracy *ab initio* methods scale as n^7 or higher. These algorithms are also well parallelized, being able to take advantage of and use hundreds, thousands, or even millions of cores efficiently. These methods can be applied to a variety of electronic structure problems, from atoms to solids.[18,19] QMC methods are routinely applied to solid-state problems in which traditional methods simply cannot be applied. Since QMC methods make use of sampling methods, there is an uncertainty associated with each calculation.

For efficient sampling, these methods require an initial reference called a trial wave function, which can be constructed from various methods, such as HF for a single-configurational trial wave function, or MCSCF for a multi-configurational trial wave function. To describe dynamic electron correlation, QMC methods make use of a Jastrow factor,[20,21] which make the trial wave function depend explicitly on interparticle separations, similar to explicitly correlated F12 methods.[22,23] A multi-determinant trial wave function can be represented mathematically in Equation 10:

$$\Psi_T(\boldsymbol{R}) = e^{J(\boldsymbol{R})} \sum_{k=1}^{N_D} c_k D_k^{\uparrow} D_k^{\downarrow} \tag{10}$$

where Ψ_T is the trial wave function, e^J is the Jastrow factor, c_k are the determinant coefficients for the multi-determinant expansion describing static correlation, and $D_k^{\uparrow}$ and $D_k^{\downarrow}$ are the spin-up and spin-down Slater determinants, respectively. Of the QMC methods, two are presented: variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC).

VMC calculates the expectation value of the Hamiltonian using Monte Carlo integration. Mathematically, it can be represented as:

$$E_{VMC} = \frac{\int \Psi_T^*(\boldsymbol{R}) H \Psi_T(\boldsymbol{R})\, d\boldsymbol{R}}{\int \Psi_T^*(\boldsymbol{R}) \Psi_T(\boldsymbol{R})\, d\boldsymbol{R}} \tag{11}$$

VMC is primarily used to optimize parameters of the trial wave function, such as the Jastrow factor, for subsequent use in the more accurate DMC. The time-dependent SE can be written mathematically as a generalized diffusion equation

$$\frac{\partial \Psi(r,\tau)}{\partial \tau} = \frac{1}{2}\nabla^2 \Psi(r,\tau) - V(r)\Psi(r,\tau) \tag{12}$$

A random walk procedure is used to simulate equation 10, in which an initial distribution of walkers is allowed to diffuse and multiply in a series of finite time steps. As the simulation continues in time and in the number of iterations, the distribution of walkers approaches a fluctuation about an average steady-state distribution corresponding to the lowest energy wave function that satisfies the time-independent SE.[24] In the long time limit, the method converges to the ground state wave function.

$$\Psi(r, \tau \to \infty) \simeq \exp[-(E_0 - E_T)\,\tau] c_0(0) \psi_0(r) \tag{13}$$

For electronic wavefunctions, since electrons are Fermions and require an antisymmetric wavefunction, DMC uses the fixed-node approximation, *i.e.*, the nodes are fixed by the trial wave function. If the trial wave function has the correct nodal structure of the exact wave function, the SE will be solved exactly. Since the nodes are not usually known, then the fixed-node DMC method gives the best wave function while constrained by the nodes of the trial wave function.

PAPER

I. PHOTOINDUCED ELECTRON TRANSFER IN DONOR-ACCEPTOR COMPLEXES OF ETHYLENE WITH MOLECULAR AND ATOMIC IODINE

Aimable Kalume,[1] Lisa George,[1] Andrew D. Powell,[2] Richard Dawes[*,2] and Scott A. Reid[*,1]

[1] *Department of Chemistry, Marquette University, Milwaukee, WI 53233*

[2] *Department of Chemistry, Missouri University of Science and Technology, Rolla, MO 65409*

ABSTRACT

Building upon our recent studies of radical addition pathways following excitation of the I_2 chromophore in the donor-acceptor complex of ethylene and I_2 ($C_2H_4\cdots I_2$), in this article we extend our studies to examine photoinduced electron transfer. Thus, irradiation into the intense charge transfer band of the complex (λ_{max} = 247 nm) gave rise to a band at 366 nm which is assigned to the bridged ethylene-I radical complex on the basis of our prior work. The formation of the radical complex is explained by a mechanism that involves rapid back electron transfer leading to I-I bond fission. Excitation into the charge transfer band of the radical complex led to regeneration of the parent complex, and the formation of the final photoproduct, *anti*– and *gauche*–1,2-diiodoethane, which confirms that the reaction proceeds ultimately by a radical addition mechanism. This finding is contrasted with our previous study of the $C_2H_4\cdots Br_2$ complex, where CT excitation led to only one product, *anti*–1,2-dibromoethane, a result explained by a single electron transfer mechanism proceeding via a bridged bromonium ion intermediate. For the I_2 complex, the breakup of the photolytically generated $I_2^{-\bullet}$ anion radical is apparently

sufficiently slow to render it uncompetitive with back electron transfer. Finally, we report a detailed computational examination of the parent and radical complexes of both bromine and iodine, using high level single– and multi–reference methods, which provide insight into the different behaviors of the charge-transfer states of the two radicals and the role of spin-orbit coupling.

Keywords: matrix isolation, neon matrix, donor-acceptor complex, ethylene-iodine.

1. INTRODUCTION

The electron donor-acceptor (EDA) complexes of halogens with π-electron donors such olefins and aromatic compounds are model systems for examining halogen bonding and the mechanism of electrophilic addition and substitution reactions.[1-5] Such complexes, which were initially observed and described by Benesi and Hildebrand in 1949,[6] and later characterized by Mulliken,[5,7,8] have been shown to be important precursors in a variety of organic electron transfer reactions,[3] and the study of ion pairs generated by excitation of EDA complexes has been used to examine many key issues in electron transfer theory.[9]

The reaction of dihalogens such as X_2 (X = F, Cl, Br, I) with the simplest π-electron donor, ethylene, is a textbook organic reaction of electrophilic addition, which is usually presumed to proceed through an EDA complex. On the basis of electron affinity, the halogen reactivity is predicted to decrease in the order F > Cl > Br > I. Recently, we investigated the mechanism of electrophilic addition in the $C_2H_4\cdots Br_2$ EDA complex by initiating photoinduced electron transfer of the complex trapped in an Argon matrix at 5 K.[2] Following excitation into the intense UV charge-transfer band, only the *anti*-conformer of the 1,2-dibromoethane product was observed, in agreement with a single

electron transfer mechanism proceeding through a bridged bromonium ion intermediate. In contrast, excitation of the Br_2 chromophore leads to Br-Br bond fission and radical addition, producing the *anti*– and *gauche*–conformers in nearly equal yield.

In another recent study, we examined radical pathways to electrophilic addition in the I_2 + ethylene system by irradiating the matrix-isolated EDA complex in the visible I_2 absorption band, leading to cleavage of the I-I bond.[10] In these steady state experiments, we observed formation of the bridged iodoethyl radical and both conformers of the 1,2-addition product, with a 2:1 preference for the *anti*-conformer. This was explained by stepwise addition via the bridged intermediate. These observations were supported by theory, which predicted that the bridged radical is the global minimum on the C_2H_4I potential energy surface (PES).

In this work we turn our attention to the photoinitiated charge transfer chemistry of the ethylene complex with molecular iodine. There are few previous reports on experimental and theoretical studies of molecular iodine complexes with ethylene,[11-13] but many studies of iodine complexes with arenes.[4,14-23] In an elegant set of experiments, Zewail and co-workers examined photoinduced electron transfer in arene (AR)$\cdots I_2$ complexes formed in the gas-phase in a supersonic beam using mass spectrometry combined with ultrafast laser spectroscopy.[17] Since in the gas-phase ion pair separation is highly endoergic, the final products arise from two channels – an ionic channel leading to $AR^{+\bullet}\cdots I^{-\bullet}$ + I, and a neutral channel producing AR$\cdots$I + I or AR + I + I.[17] The latter can be thought as arising from back electron transfer (BET), which produces an excited I_2 molecule at the same energy as the CT state. At this energy the nascent I_2 molecule is found on a repulsive electronic surface, leading to rapid bond cleavage. The energy

release into these channels was also characterized by Young and co-workers.[19,20] In the solution phase, a variety of ultrafast studies of arene···I_2 complexes have been performed.[16,17,21-24] These have revealed the rapid (sub ps) formation of an arene···I complex, which is presumed to arise from a similar mechanism (i.e., BET followed by I-I bond fission) as elucidated in the gas-phase studies.

With these results in mind, here we report experimental and computational studies of photoinduced electron transfer in the C_2H_4···I_2 donor-acceptor complex, trapped in Ne matrices at ~ 4 K. We also report the first studies, to our knowledge, of the photoinduced electron transfer in this complex following excitation into the intense charge-transfer band in the UV region. Our experimental results are supported by Density Functional Theory (DFT) calculations and high level single– and multi–reference calculations, which have examined in detail the influence of spin-orbit coupling on the binding energies of molecular and radical complexes.

2. EXPERIMENTAL AND COMPUTATIONAL METHODS

The experimental apparatus has been described in detail in earlier publications.[2,25,26] Briefly, a mixture of ethylene:Ne (~1:500) was prepared in a 0.5 L mixing tank using standard manometric methods. This was passed over iodine crystals held in a home-made sample holder that was heated to 5-10 degrees above room temperature, and sprayed onto a cold window held at ~ 5 K using the pulsed deposition method with a solenoid activated pulsed valve (Parker-Hannifan, General Valve Division, Iota-1). Typical deposition conditions were: 1 ms pulse duration, 5 Hz repetition rate, 1-2 h deposition time, 1 bar backing pressure.

Following deposition, IR spectra were obtained with an FTIR spectrometer (ThermoNicolet 6700) equipped with a DTGS detector, which was purged at a flow rate of 20 L/min using a purge gas generator (Parker-Balston 75-52A). IR spectra were recorded at 1 cm^{-1} resolution and typically averaged over 128 scans. UV-Visible spectra (200-1100 nm) were obtained using an Agilent diode array spectrophotometer, with a typical integration time of 1 s. All spectra were referenced to the cold sample window, and subsequently transferred to a spreadsheet and analysis program (Origin 9.0) for workup. Photolysis of the complex was initiated by irradiating the cold window with laser light at 240 nm, generated from the frequency doubled output of a dye laser system (Lambda-Physik Scanmate 2E) operating on Coumarin 480 dye, pumped by the third harmonic (355 nm) of a Nd:YAG laser (Continuum NY-61). The photolysis beam was expanded using a 4:1 beam expander to fill the cold window and avoid damage to the KBr windows. Typical irradiation times were 1-2 h. In some experiments, a second photolysis step was performed with laser light at 355 nm, generated from the third harmonic of a pulsed Nd:YAG laser (Continuum Minilite).

Initial geometry optimizations were performed using Density Functional Theory (M06 and M06-2X functionals) methods with Peterson's aug-cc-pVTZ-pp basis set for Iodine and an aug-cc-pVTZ basis set for all other atoms, as implemented in the Gaussian09 program.{Frisch, 2010 #34} Electronic absorptions and oscillator strengths were calculated using time-dependent DFT, with the CAM-B3LYP and M06 functionals. Subsequently, multistate MRCI calculations were performed using Molpro[27] to examine the low-lying singlet and triplet states of the ethylene-I_2 complex as well as their correlations upon dissociation to form the bridged radical ethylene-I· species, and the

effects of spin-orbit (SO) coupling. These calculations used Peterson's new triple-zeta post-d F12 basis sets for Iodine,[28] which has 25 explicit electrons surrounding a relativistic pseudo-potential representing 28 core-electrons. Calculations of the structure and harmonic frequency of I_2 as well as the splitting due to spin-orbit coupling for atomic iodine are in close agreement with experiment (see *Supporting Information*). The CASSCF level is at least semi-quantitative and is easily affordable for scanning multiple states of the complexes along the dissociation coordinate. We also present some limited MRCI-F12 results, correlating as many electrons as was feasible.

3. RESULTS AND DISCUSSION

Figure 1 displays UV/Visible spectrum of C_2H_4:I_2:Ne (~1:2:500) sample at ~5 K, and a stick spectrum representing the TDDFT prediction of the spectrum of ethylene-iodine complex at the (TD)M06/aug-cc-pVTZ level. When ethylene and iodine were co-deposited in neon matrix an intense band appeared at 247 nm, which is in excellent agreement with theoretical predictions (Table 1), being essentially bracketed by the TDCAM-B3LYP and TDM06 predictions.

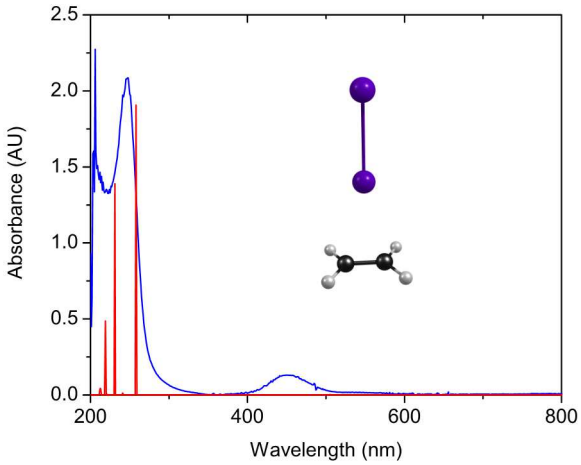

Figure 1. The UV-Visible spectrum (blue) of the Ethylene/I_2/Ne (~2:1:500), with the predicted spectrum (red, shown as stick spectrum) of the singlet states of the ethylene-I_2 complex at the TDM06/aug-cc-pVTZ-pp level.

Table 1. Predicted and observed electronic absorptions for C_2H_4 complexes. [a] With aug-cc-pVTZ-pp basis set on optimized M062x/aug-cc-pVTZ-pp geometry. [b] In Ar matrix.[11] The excited states of the molecular complex are of singlet spin multiplicity, while those of the radical complex are of doublet spin multiplicity.

Complex	**Excited State**	**Predicted vertical excitation λ in nm (oscillator strength, *f*)**			**λ_{max} (*f*)**	**λ_{max} (nm)**
		CAM-B3LYP[a]	**M06[a]**	**M06–2X[a]**	**Ne**	**Other**
$C_2H_4\cdots I_2$	1	497.3(0.0009)	505.0(0.0008)	494.7(0.0009)	…	…
	2	494.7(0.0005)	502.6(0.0005)	492.7(0.0006)	…	…
	3	296.9(0.0001)	310.6(0.0001)	293.7(0.0001)	…	…
	4	296.3(0.0003)	310.1(0.0002)	293.4(0.0004)	…	…
	5	237.4(0.7591)	257.6(0.4240)	235.7(0.8043)	247	248.5[b]
	6	196.2(0.0001)	240.5(0.0013)	194.0(0.0008)	…	…
	7	195.8(0.0004)	240.4(0.0030)	193.8(0.0000)	…	…
	8	194.5(0.0000)	238.9(0.0000)	192.8(0.0000)	…	…
	9	189.5(0.2660)	231.1(0.3088)	188.0(0.1692)	…	…
	10	177.6(0.2248)	222.8(0.0000)	182.8(0.2877)	…	…
$C_2H_4\cdots I\bullet$	1	2434.2(0.0000)	2525.6(0.0000)	1906.0(0.0001)	…	…
	2	2337.5(0.0001)	2460.1(0.0001)	1900.5(0.0001)	…	…
	3	356.4(0.2036)	383.5(0.1724)	300.6(0.1671)	366	…
	4	306.9(0.0008)	304.1(0.0012)	276.2(0.0006)	…	…
	5	229.2(0.0000)	264.1(0.0000)	205.6(0.0000)	…	…
	6	198.3(0.0013)	238.2(0.0037)	193.8(0.0000)	…	…
	7	196.7(0.0000)	237.8(0.0000)	193.5(0.0157)	…	…
	8	193.7(0.0000)	235.7(0.0075)	190.0(0.0247)	…	…
	9	190.8(0.0256)	235.6(0.0108)	189.8(0.0447)	…	…
	10	190.0(0.0178)	231.3(0.0000)	188.6(0.0402)	…	…

The position of the charge-transfer band in the Ne matrix is similar to previous results in Ar and nitrogen matrices (248.5 nm in Argon and 246.5 nm in N_2).[11] The highest-level CCSD(T)-F12b/VTZ-F12 calculation predicts a binding energy of 17.47 kJ/mol for the C_2H_4:I_2 complex (slightly more than the value of 15.96 kJ/mol predicted for the C_2H_4:Br_2 complex at the same level of theory).

Table 2. Predicted vertical excitation with and without spin-orbit coupling for the bridged iodoethyl racdical. [a] Fine structure levels occur in precisely degenerate pairs.

Excited state	**Predicted vertical excitation λ in nm**	
	Without SO	**With SO[a]**
1	6573.6	9158.1
2	6355.2	1247.9
3	340.5	324.6
4	283.1	284.2
5	275.4	278.3
6	274.2	277.8
7	272.4	276.2
8	271.8	275.5
9	263.3	262.1
10	184.0	232.3

The formation of the complex is also evident in the IR spectra, by the appearance of bands that are IR inactive in ethylene due to symmetry lowering (Figure S1 in the supporting information). The IR inactive ethylene bands ν_2 (C-C stretch) and ν_3 (CH_2 deformation) appear at 1342 and 1613 cm^{-1}, respectively, upon formation of the complex, while the infrared active ethylene bands display only small shifts upon complex formation.[10] Under the assumption that the IR and UV/Visible spectra sample the same region of the matrix, the integrated IR and UV/Vis intensities can be combined with calculated IR intensities to estimate the oscillator strength of the UV/Vis transitions. The integrated IR absorbance of a given feature was divided by the calculated intensity (M06-

2X/aug-cc-pVTZ-pp, in km/mol) to derive a column density in the matrix, and an average value was obtained over the observed IR absorptions. The oscillator strength of a given electronic (UV/Visible) band was then obtained as:[31]

$$f = \frac{\int A_{UV}(\upsilon)\, d\upsilon}{N_{IR}} \times \left(1.87 \times 10^{-7}\ \text{mol/km}\right) \quad (1)$$

where N_{IR} is the column density derived from the IR measurements and the numerator represents the integrated ultraviolet absorbance (over cm^{-1}). Using this approach, the derived oscillator strength of the CT band is ~ 0.46 (ε_{max} = 12,900 M^{-1} cm^{-1}), similar to the TD M06 prediction (Table 1).

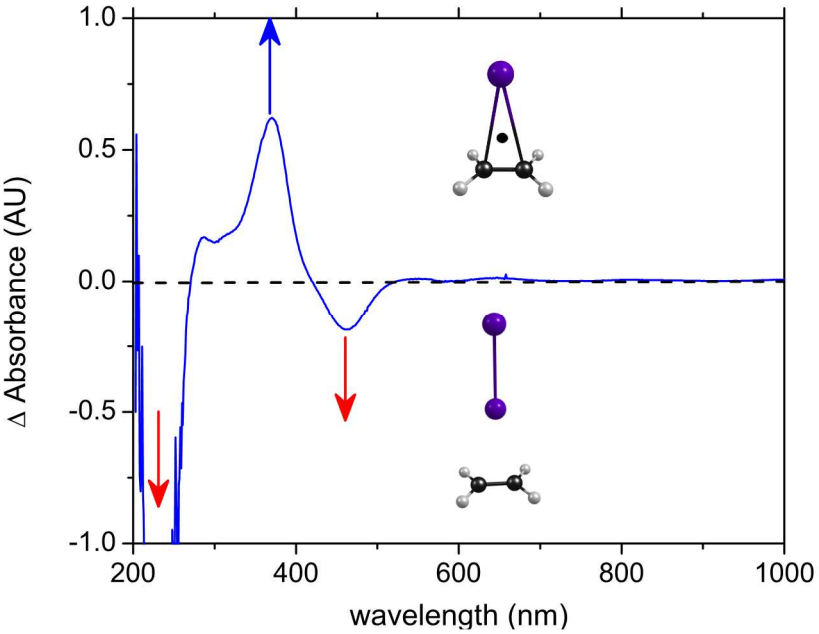

Figure 2. Difference UV-Visible spectrum obtained following 240 nm irradiation of an ethylene:I_2:Ne (~2:1:500) matrix. The disappearance of features assigned to the ethylene-I_2 complex coincides with the rise of a feature at 366 nm which is assigned to the bridged iodoethyl radical.

The CT photochemistry of the complex was probed by 240 nm laser irradiation of the matrix. Figure 2 demonstrates that irradiation leads to the loss of the $C_2H_4\cdots I_2$ band at 247 nm and the appearance of an intense band at 366 nm, which is readily assigned to the bridged $C_2H_4\cdots I$ complex on the basis of our earlier work. The observed red-shift in this absorption is similar to that observed for the corresponding complexes with benzene ($Bz\cdots I_2$, λ_{max} = 295 nm; $Bz\cdots I$, λ_{max} = 430 nm), and the position of this band is in excellent agreement with TDDFT predictions, particularly when the M06 functional is used (Table 1). The experimentally determined oscillator strength determined using equation (1) is also in good agreement with theory (Table 1). The formation of the radical complex is also evident in the IR spectra (Figure S1, supporting information). For example, considering the 1613 cm^{-1} absorption, difference-spectra of the matrix obtained pre- and post-photolysis reveal a decrease in intensity of this feature and growth of a feature at 1589 cm^{-1}.[10]

Previous studies of $arene\cdots I_2$ complexes in the gas-phase and solution have shown that CT excitation followed by rapid back electron transfer leads to formation of a neutral I_2 molecule on an excited repulsive potential energy surface, resulting in rapid fission of the I-I bond.[4,14-18,22-24,32,33] Zewail and co-workers identified a second "harpoon" like mechanism that led to a smaller translational energy release,[17] and in the gas-phase accounts for some 30-40% of products.[19,20] In the matrix, cleavage of the I-I bond following BET leads to formation of the separated radical pair C_2H_4I and I. That the same bridged C_2H_4I radical is observed following excitation into both the CT and Visible bands of the $C_2H_4\cdots I_2$ complex speaks to the fast BET that must follow CT excitation in this system.

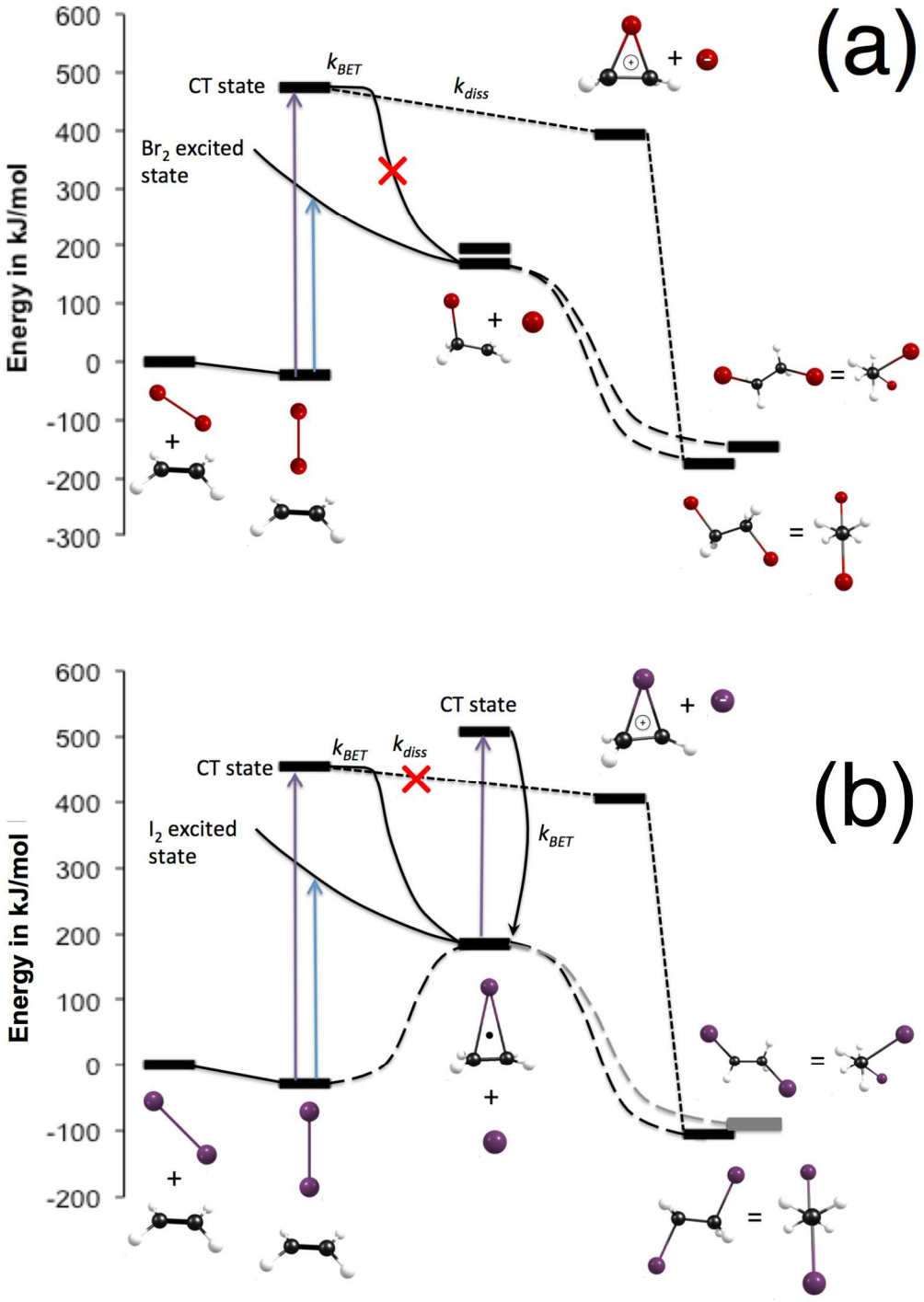

Figure 3. Schematic of the photochemical pathways of the $C_2H_4\cdots Br_2$ (a) and $C_2H_4\cdots I_2$ (b) complexes.

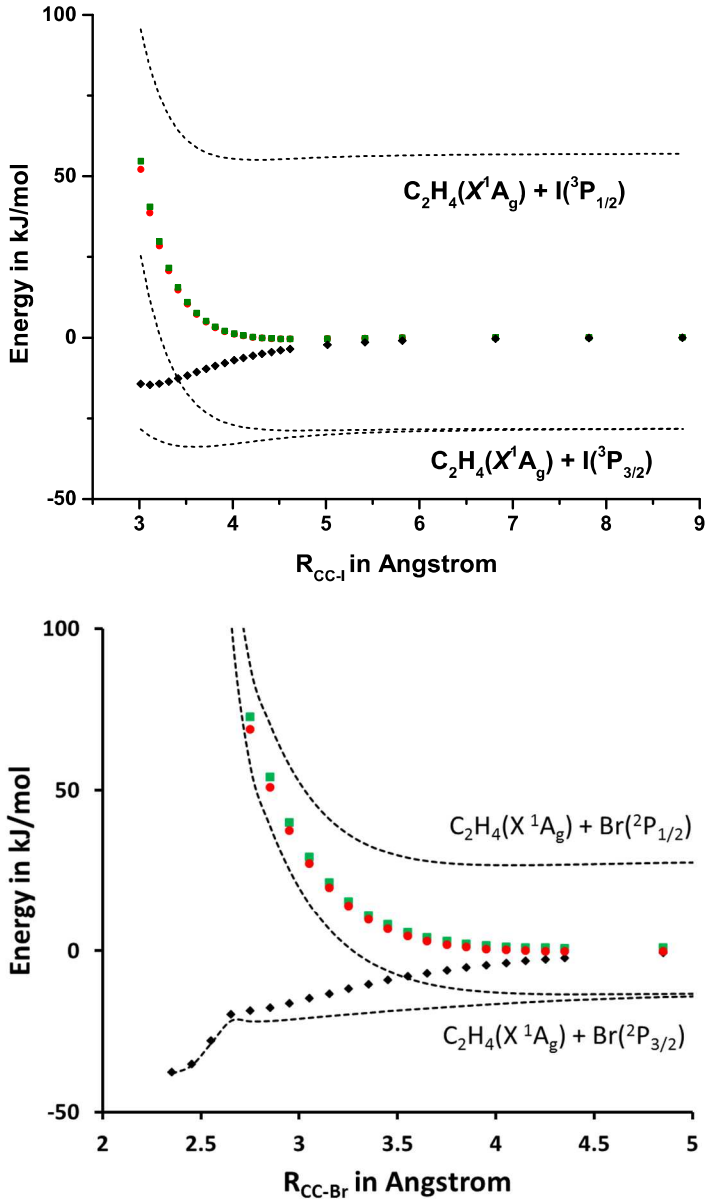

Figure 4. Plot of A_1, B_1 and B_2 states, scanning the distance from the C-C bond midpoint (X) to I (or Br), in C_2H_4---I· (upper panel) and C_2H_4---Br· (lower panel) with and without the effects of SO-coupling. The Davidson-corrected MRCI(Q)-F12 method was used to compute the energies and couplings. The dashed lines show three pairs of fine-structure states correlated with the I($^2P_{1/2}$) and I($^2P_{3/2}$) atomic states compared with the results obtained without SO coupling (solid symbols).

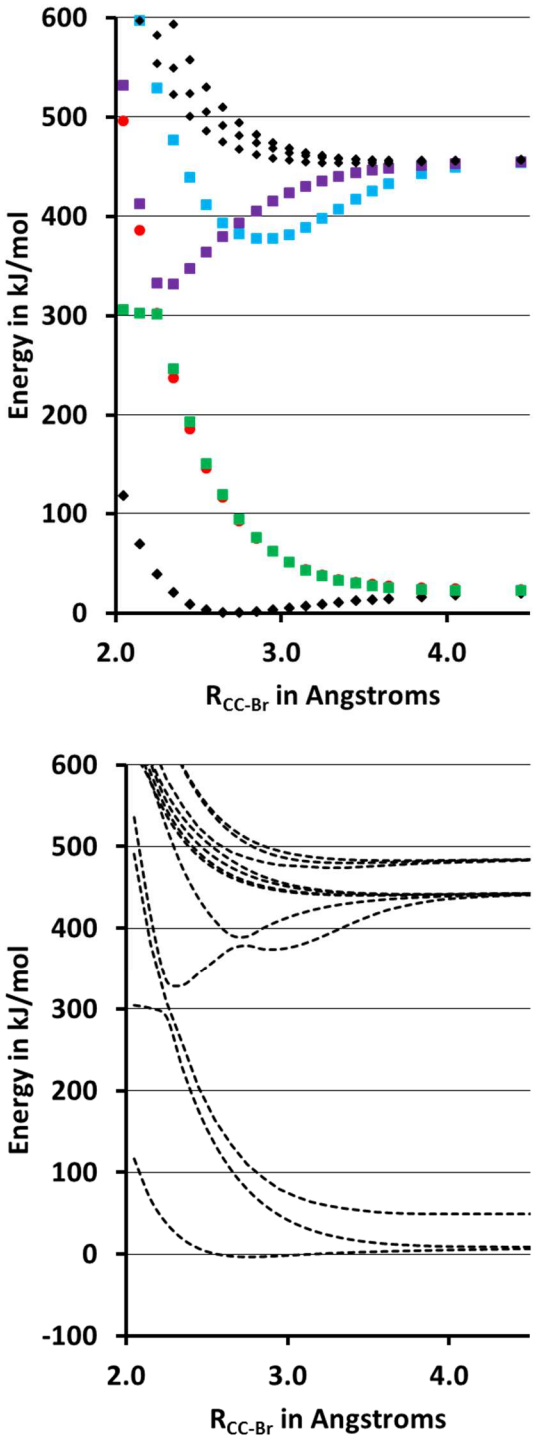

Figure 5. Plot of excited states scanning the distance from the C-C bond midpoint (X) to I in C_2H_4---I· at the MRCI-F12 level. (upper) without SO coupling. (lower) pairs of fine-structure states obtained by including the effects of SO-coupling.

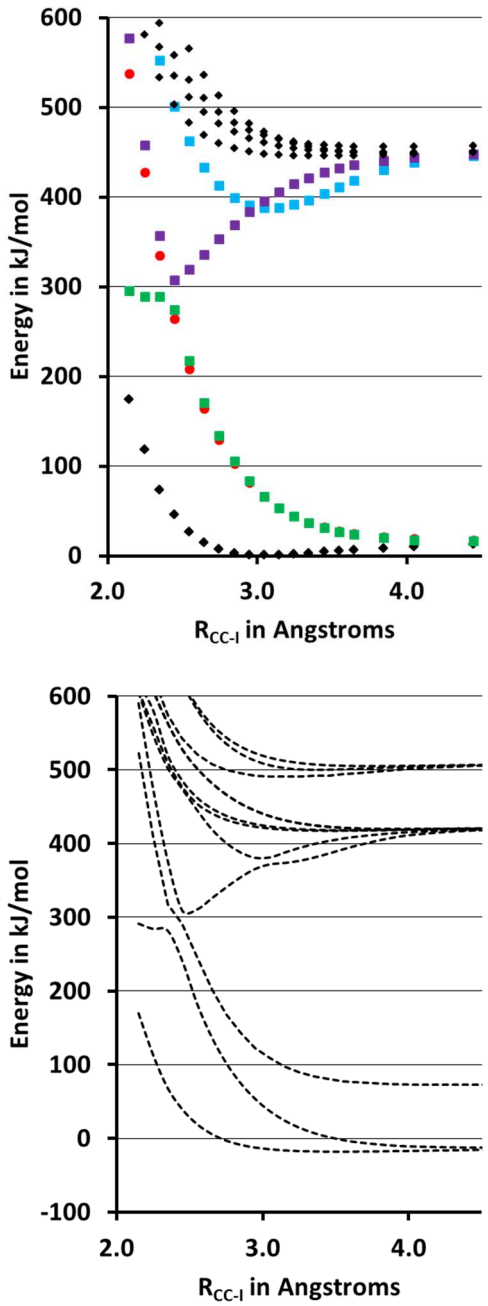

Figure 6. Plot of excited states scanning the distance from the C-C bond midpoint (X) to Br in C_2H_4---Br· at the MRCI-F12 level. A direct C2V approach is followed (not the relaxed path shown in Figure 4. (upper) without SO coupling. (lower) pairs of fine-structure states obtained by including the effects of SO-coupling.

Our observation of the radical complex suggests that the second iodine atom has escaped the primary cage, avoiding geminate recombination. However, upon excitation

of CT band of the radical complex, regeneration of the parent $C_2H_4\cdots I_2$ complex is observed, as are the two conformers of the 1,2-addition product. These experiments probed the dynamics of the $C_2H_4\cdots I$ complex by performing a second irradiation at 355 nm following a first irradiation at 240 nm. The UV-Visible difference spectrum shows a loss of the 366 nm band of the $C_2H_4\cdots I$ complex, with a concomitant increase in bands of the $C_2H_4\cdots I_2$ complex, while the IR spectrum shows the appearance of 1,2-diiodoethane. Following electron transfer, the $C_2H_4^+\cdots I^-$ intermediate can decay via BET to ethylene and a secondary iodine atom, which can recombine with the primary atom to form I_2. Alternatively, a sequential radical addition of the two iodine atoms to ethylene can yield *anti*– and *gauche*–1,2-diiodoethane.

Barbara and co-workers examined the dynamics following CT excitation in related arene···Br atom complexes in various solvents using ultrafast transient absorbance spectroscopy over a range of probe wavelengths.[34,35] Typically, the observed BET kinetics exhibited multiexponential behavior, with a fast component (ca. 1 ps), a slow component (ca. 0.2–1 ns), and intermediate components. The multiexponential behavior was attributed to a distribution of arene+/Br- geometries in the initial ion pair state, which was presumed to involve specific, geometry dependent electronic interactions between donor and acceptor.[35]

It is instructive to compare the CT photochemistry of $C_2H_4\cdots I_2$ with our previous study of the $C_2H_4\cdots Br_2$ complex.[2] The combined insights are summarized in Figure 3. For the Br_2 complex, CT excitation leads to only one product, *anti*–1,2-dibromoethane, which is explained by a single electron transfer mechanism proceeding via a bridged

bromonium ion intermediate (Figure 3a). For this ion to form, the breakup of the $Br_2^{-\bullet}$ anion radical (k_{diss} in Figure 3a) must be sufficiently fast to compete with BET. In contrast, excitation of the visible Br_2 centered absorption leads through a classical haloaklyl intermediate to both stereoisomers of the 1,2-addition product, in nearly equal yield, consistent with a radical addition mechanism. For the I_2 complex, the lower reactivity of $I_2^{-\bullet}$ favors BET, and the fragmentation therefore occurs on a repulsive neutral potential surface, leading to formation of a bridged haloaklyl radical (Figure 3(b)).[36] Thus, in this case the product yield is similar to that obtained through excitation of the visible I_2 centered absorption. This work illustrates that, following CT excitation, the competition between BET and fragmentation of the radical anion is important in controlling the subsequent chemistry.

In order to examine the two complexes in more detail and the role of spin-orbit coupling, calculations were performed at the UCCSD(T)-F12b, SA-CASSCF, and MRCI-F12 levels. Where applicable (i.e., for the lowest energy state of a particular symmetry and spin in regions where multiple configurations are not important), the UCCSD(T)-F12b method provides highly accurate benchmarks, benefitting from inclusion of (perturbative) triples in the correlation treatment. The SA-CASSCF provides at least a semi-quantitative description of the behavior of multiple states and the effects of SO coupling, and MRCI-F12 (to the extent it is affordable) provides greater accuracy to the multistate calculations. The iodine complex structures were optimized using UCCSD(T)-F12b, yielding an inter-iodine distance in C_2H_4---I_2 of $r_{I\text{-}I}$ = 2.702 Å, only slightly longer than in the I_2 diatomic. The distances from the C-C bond midpoint to the I-atom in C_2H_4---$I\cdot$ and the closer I-atom in C_2H_4---I_2 were 2.98 Å and 3.13 Å respectively. This distance

is sensitive to the correlation treatment. MRCI calculations with a full-valence active space and all electrons correlated are prohibitively expensive, and successive reductions in the number of orbitals included in the correlation procedure resulted in correspondingly longer optimized distances. The geometric parameters of ethylene are insensitive to the presence of iodine atoms in these complexes. Values for the C-C bond distance of 1.334 Å, 1.343 Å, and 1.338 Å were recorded for C_2H_4, C_2H_4---I· and C_2H_4---I_2 respectively. Other geometric parameters such as C-H bonds and CCH angles varied even less. The hydrogen atoms do bend very slightly out-of-plane (away from the iodine atom(s)). For the bromine complex C_2H_4---Br_2 the behavior is similar to that of the iodine system since the calculated inter-bromine distance of 2.311 Å is only slightly longer than in the Br_2 diatomic and again the structure of ethylene is not strongly perturbed. The distance to the closer Br atom is 2.966 Å. Whereas the ground state of the weakly bound C_2H_4---I· complex is bridged, the corresponding bromine system C_2H_4---Br· forms a more strongly bound classical radical with the bromine atom intimately bonded to one carbon.

A relaxed scan was performed beginning with the C_2H_4---I_2 complex (binding energy of 17.47 kcal/mol at the CCSD(T)-F12b/VTZ-F12 level), increasing the $r_{I\text{-}I}$ distance (asymptotically reaching the C_2H_4---I· complex). Additional details and results can be found in SI. The scan using the UCCSD(T)-F12b method was only used to obtain a series of geometries at which to plot energies obtained by the state-averaged CASSCF (SA-CASSCF) method. As the I_2 fragment is pulled apart the optimized distance from the C-C bond midpoint to the near I-atom initially decreases to a minimum of 2.803 Å, corresponding to an $r_{I\text{-}I}$ distance of 3.6 Å, before recovering toward the asymptotic value of 2.98 Å for the C_2H_4---I· structure.

Without considering SO-coupling, a single iodine atom associated with ethylene to form the C_2H_4---I· complex breaks the degeneracy of the A_1, B_1, and B_2 states. The stabilization of the $^1A_1$ state likely comes from the electrostatic interactions between the quadrupole moments of I• and ethylene molecule, with the positive side of the I• quadrupole due to the electron hole interacting with the negative side of the ethylene quadrupole due to the π-cloud. In the $^1B_1$ and $^1B_2$ states, the I• atom faces the ethylene π-cloud with a doubly-occupied *p*-orbital, and these states are thus highly repulsive. Indeed, the $^1B_1$ and $^1B_2$ states appear 97.1 and 101.7 kJ/mol above the ground $^1A_1$ respectively at the SA-CASSCF level. More quantitative descriptions obtained with UCCSD(T)-F12 and MRCI-F12 produce gaps to the B_1 and B_2 states of about 65-70 kJ/mol (discussed in more detail below).

Considering the C_2H_4---I· complex, SO-coupling is predicted to significantly impact the states and their splittings, even shifting the equilibrium distance to the I-atom by ~0.6 Angstroms. The ground states of C_2H_4 and I-atom combine to yield three states (A_1, B_1 and B_2), the lowest (A_1) having its minimum (without SO-coupling) at a distance of 2.98 Angstroms by the UCCSD(T)-F12b method as discussed above. Since the other two states are each the lowest states of different symmetries, they can also be calculated at the UCCSD(T)-F12b level. At the equilibrium distance without SO-coupling (2.98 Å), the gaps to the B_1 and B_2 states are 68.3 and 71.2 kJ/mol at the UCCSD(T)-F12b level, and the well depth on the ground A_1 state is 20.3 kJ/mol. Dynamic electron correlation is quite important, as MRCI-F12 produces similar gaps to the B_1 and B_2 states of 66.6 and 69.1 kJ/mol, respectively, and a well depth on the A_1 state of 14.6 kJ/mol, but the SA-CASSCF calculation produces a much shallower well (1.7 kJ/mol) at larger separation

and is significantly repulsive where the UCCSD(T)-F12b and MRCI-F12 methods have their minima. Figure 4 plots the three states A_1, B_1 and B_2 at the MRCI-F12 level as a function of distance from the C-C midpoint. Also shown are the three pairs of fine-structure levels produced by including the effects of SO-coupling. Remarkably, the minimum on the lowest pair of fine-structure levels is shifted outward by more than 0.6 Å, relative to the result obtained without considering SO-coupling.

The binding in the C_2H_4---Br· complex is quite different. At large distances from the C-C midpoint the behavior is similar to that of the C_2H_4---I· system with the Br-atom preferring to associate directly above (C_{2V} symmetry) a negligibly perturbed ethylene. Beginning at a distance of ~2.75 Å the Br atom prefers to bond more intimately with one C-atom (lowering symmetry to C_S) and ethylene relaxes rapidly toward the classical radical structure ($r_{C\text{-}Br}$ = 2.014 Å). This transition appears as a shoulder in the energy along the minimum energy path (MEP) shown in Figure 4. SO-coupling becomes negligible at the shortest bond distance and thus in contrast to the iodine system does not significantly perturb the energy or structure of the radical. Additional plots showing the relaxation of the angle and C-C distance along the MEP are included in SI.

Plots of the excited states of C_2H_4---I· and C_2H_4---Br· (Figures 5 and 6) provide some insight into the different behaviors of the charge-transfer states of the two radicals. Figure 5 plots the excited states along the direct (C_{2V}) approach of a Br atom. Thus the zero of energy is the shoulder of the curve in Figure 4 just before the MEP deviates sharply to form the classical radical bound to one C-atom. The lower panel shows the effects of SO-coupling and plots the fine-structure pairs. In Figure 6 corresponding plots are shown for C_2H_4---I·. Experimentally irradiation at 355 nm excites C_2H_4---I· into the

CT state. This corresponds to an energy of ~398 kJ/mol which is close to the gap to a pair of fine-structure levels above the minimum on the ground fine-structure state (at 3.6 Å, see Figure 4). It can be seen that SO-coupling affects the stability of the state providing a path to the repulsive $C_2H_4(X\ ^1A_g) + I(^2P_{1/2})$ component of the ground state.

An effort was made to make more quantitative predictions of vertical excitation energies for the bridged C_2H_4---I· radical at the MRCI-F12 level with and without the effects of spin-orbit coupling. Calculations were performed using the structure corresponding to the minimum on the lowest fine-structure levels shown in Figure 4 (with a distance from the C-C bond midpoint to the I-atom of 3.6 Å as discussed above). Some details can be found in SI. In brief, MRCI-F12 calculations using a Rydberg extended basis set, and the largest affordable active space were employed. A total of 11 doublet (4 A_1, 3 B_1, 3 B_2, and 1 A_2) and 5 quartet (1 A_1, 1 B_1, 2 B_2, and 1 A_2) state energies were computed including an applied *rotated* Davidson correction. The scans in Figures 5 and 6 employed a less costly reduced active space and fewer states (the lowest 6 doublet and 3 quartet states). Energies obtained this way were combined with SO-couplings computed without excitations (at the CASSCF level), and these are listed in Table 2. They are only in rough accord with experiment, which suggests that the agreement of experiment and TDDFT predictions (Table 1) for the radical complex is probably fortuitous.

4. CONCLUSIONS

The donor-acceptor complex of ethylene and I_2 ($C_2H_4\cdots I_2$) was isolated in a neon matrix at 5 K and characterized by infrared and electronic spectroscopy, supported by *ab initio* and Density Functional Theory (DFT) calculations. Irradiation into the intense CT band (λ_{max} = 247 nm, f=0.46) of the complex gave rise to a band at 366 nm that is assigned to the bridged $C_2H_4\cdots I\bullet$ radical complex on the basis of (TD)DFT calculations. Following CT excitation, the formation of the radical complex ($C_2H_4\cdots I\bullet$) is explained by rapid back electron transfer leading to I-I bond fission, and this is supported by observation of the same complex following excitation of the visible absorption band (λ_{max} = 451 nm) of the $C_2H_4\cdots I_2$ complex, which is centered on the I_2 chromophore. The radical complex is the only photoproduct observed following either UV or Visible irradiation.

To probe the spectroscopy and photochemistry of the radical complex, irradiation into the charge transfer band was performed, at 355 nm. In addition to the $C_2H_4\cdots I_2$ complex, IR bands assigned to *anti*– and *gauche*–1,2-diiodoethane are observed, which confirms that the reaction proceeds by radical addition, rather than through an iodonium ion intermediate. This stands in contrast to our previous study of the $C_2H_4\cdots Br_2$ complex, where CT excitation led to only one product, *anti*–1,2-dibromoethane, a result explained by a single electron transfer mechanism proceeding via a bridged bromonium ion intermediate. In this case, the breakup of the $Br_2^{-\bullet}$ anion radical generated following CT was sufficiently fast to compete with BET; however, for the I_2 complex, the slower fragmentation of $I_2^{-\bullet}$ favors BET and subsequent I-I bond fission on a repulsive neutral potential energy surface.

The observed UV-Visible absorptions and associated oscillator strengths, and IR bands of the $C_2H_4\cdots I_2$ and $C_2H_4\cdots I\bullet$ complexes are largely in excellent agreement with (TD)DFT predictions. For the radical complex, this is possibly fortuitous, as high level multireference calculations reveal the important role of spin-orbit coupling in this system, which significantly reduces the binding energy of the ground state complex and leads to a rich set of crossings and avoided crossings.

In future studies, it would be particularly revealing to probe the short-time dynamics following photoinduced electron transfer of this prototypical complex in both the gas-phase and condensed phases.

ACKNOWLEDGEMENTS

Support of the research by the National Science Foundation (CHE-1057951 to S.R.) and (CHE-1300945 to R.D.) the donors of the Petroleum Research Fund of the American Chemical Society (PRF 48740-ND6) are gratefully acknowledged. A.K. was supported in part by an Eisch fellowship, and gratefully acknowledges the generosity of Dr. John J. Eisch. This research was also funded in part by National Science Foundation awards OCI-0923037 "MRI: Acquisition of a Parallel Computing Cluster and Storage for the Marquette University Grid (MUGrid)" and CBET-0521602 "Acquisition of a Linux Cluster to Support College-Wide Research & Teaching Activities." The authors thank Prof. Rajendra Rathore and Prof. Qadir Timerghazin for many useful discussions.

REFERENCES

(1.) R. Bianchini, C. Chiappe, D. Lenoir, P. Lammen, R. Herges, and J. Grunenberg, *Angewandte Chemie, International Edition in English* 36, 1284 (1997).

(2.) L. George, L. Wittmann, A. Kalume, and S. A. Reid,*J. Phys. Chem. Lett*. 1, 2618 (2010).

(3.) D. Sun, S. V. Rosokha, J. K. Kochi, *J. Am. Chem. Soc.* 126, 1388 (2004).

(4.) P. Y. Cheng, D. Zhong, and A. H. Zewail, *Chem. Phys. Lett.* 242, 369 (1995).

(5.) R. S. Mulliken and W. B. Pearson, *Molecular Complexes: A Lecture and Reprint Volume*, (Wiley Interscience,1969).

(6.) H.A. Benesi and J. H. Hildebrand, *J. Am. Chem. Soc. A* 71, 2703 (1949).

(7.) R. S. Mulliken, *J. Am. Chem. Soc.* 72, 600 (1950).

(8.) R. S. Mulliken, *J. Am. Chem. Soc.* 74, 811 (1952).

(9.) P. F. Barbara, T. J. Meyer, and M. A. Ratner, *J. Phys. Chem.* 100, 13148 (1996).

(10.) L. George, A. Kalume, and S. A. Reid, *Chem. Phys. Lett.* 554, 86 (2012).

(11.) L. Fredin,*Chemica Scripta* 4, 97 (1973).

(12.) I. Jano, *Theor Chim Acta* 66, 341 (1985).

(13.) N. Sogoshi, T. Wakabayashi, T. Momose, and T. Shida, *J. Phys. Chem. A* 101, 522 (1997).

(14.) D. P. Zhong, T. M. Bernhardt, and A. H. Zewail, *J. Phys. Chem. A* 103, 10093 (1999).

(15.) D. P. Zhong, and A. H. Zewail, *J. Phys. Chem. A* 102, 4031 (1998).

(16.) J. T. Su and A. H. Zewail, *J. Phys. Chem. A* 102, 4082 (1998).

(17.) P. Y. Cheng, D. Zhong, and A. H. Zewail, *J. Chem. Phys.* 105, 6216 (1995).

(18.) P. Y. Cheng, D. Zhong, and A. H. Zewail, *J. Phys. Chem.* 99, 15733 (1995).

(19.) G. DeBoer, J. W. Burnett, and M. A. Young, *Chem. Phys. Lett.* 259, 368 (1996).

(20.) G. DeBoer, J. W. Burnett, A. Fujimoto, and M. A. Young, *J. Phys. Chem.* 100, 14882 (1996).

(21.) H. J. Liu, S. H. Pullen, L. A. Walker, and R. J. Sension, *J. Chem. Phys.* 108, 4992 (1998).

(22.) L. A. Walker, S. Pullen, B. Donovan, and R. J. Sension, *Chem. Phys. Lett.* 242, 177 (1995).

(23.) S. Pullen, L. A. Walker, and R. J. Sension, *J. Chem. Phys.* 103, 7877 (1995).

(24.) E. Lenderink, K. Duppen, and D. A. Wiersma, *Chem. Phys. Lett.* 211, 503 (1993).

(25.) L. George, A. Kalume, P. Z. El-Khoury, A. Tarnovsky, and S. A. Reid, *J. Chem. Phys.* 132, 084503 (2010).

(26.) A. Kalume, L. George, P. Z. El-Khoury, A. N. Tarnovsky, and S. A. Reid, *J. Phys. Chem. A* 114, 9919 (2010).

(27.) H.-J. Werner, P. J. Knowles, and G. Knizia, MOLPRO ed. 2012, a package of ab initio programs. see http://www.molpro.net**.**

(28.) K. A. Peterson and J. G. Hill, *Theor. Chem. Acc.* 133, 1-12 (2014).

(29.) D. E. Kelleher, W. C. Martin, W. L. Wiese, J. Sugar, J. R. Fuhr, K. Olsen, A. Musgrove, P. J. Mohr, J. Reader, G. R. Dalton, *Phys. Scr.* T83, 158 (1999).

(30.) P. J. Linstrom and W. G. Mallard, *J. Chem. Eng. Data* 46, 1059 (2001).

(31.) H. G. Kjaergaard, T. W. Robinson, K. A. Brooking, *J. Phys. Chem. A* 104, 11297 (2000).

(32.) R. M. Whitnell, K. R. Wilson, Y. J. Yan, A. H. Zewail, *J. Mo.l Liq.* 61, 153 (1994).

(33.) E. M. Abdalla, A. A. A. A. Boraei, and M. R. Mahmoud, *Can. J. Appl. Spectrosc.* 39, 123 (1994).

(34.) A. Hormann, W. Jarzeba, P. F. Barbara, *J. Phys. Chem.* 99, 2006 (1995).

(35.) R. E. Schlief, W. Jarzeba, K. A. M. Thakur, J. C. Alfano, A. E. Johnson, and P. F. Barbara, *J. Mol. Liq.* 60, 201 (1994).

(36.) The rate of dissociative electron attachment is governed by the overlap of the repulsive ion excited states with the neutral vibrational wavefunction, {see, E. C. M. Chen and W. E. Wentworth, *J. Phys. Chem.* 89, 4099-4105 (1985).

(37.) B. Lasorne, M. A. Robb, H. D. Meyer, and F. Gatti, *Chem. Phys.* 377, 30 (2010).

(38.) K. B. Wiberg, C. M. Hadad, J. B. Foresman, and W. A. Chupka, *J. Phys. Chem.* 96, 10756 (1992).

II. CALCULATING POTENTIAL ENERGY CURVES WITH FIXED-NODE DIFFUSION MONTE CARLO: CO AND N_2

Andrew D. Powell and Richard Dawes

Department of Chemistry, Missouri University of Science and Technology, Rolla, MO 65409

ABSTRACT

This study reports on the prospect for routine use of Quantum Monte Carlo (QMC) for the electronic structure problem, applying fixed-node Diffusion Monte Carlo (DMC) to generate highly-accurate Born-Oppenheimer potential energy curves (PECs) for small molecular systems. The singlet ground electronic states of CO and N_2 were used as test cases. The PECs obtained by DMC employing multiconfigurational trial wavefunctions were compared with those obtained by conventional high-accuracy electronic structure methods such as multireference configuration interaction (MRCI) and/or the best available empirical spectroscopic curves. The goal was to test whether a straightforward procedure using available QMC codes could be applied robustly and reliably. Results obtained with DMC codes were found to be in close agreement with the benchmark PECs and the n^3 scaling with the number of electrons (compared with n^7 or worse for conventional high-accuracy quantum chemistry) could be advantageous depending on the system size. Due to a large pre-factor in the scaling, for the small systems tested here it is currently still much more computationally intensive to compute PECs with QMC. Nevertheless, QMC algorithms are particularly well-suited to large-scale parallelization and are therefore likely to become more relevant for future massively-parallel hardware architectures.

Keywords: Potential Energy Curves, Quantum Monte Carlo, DMC

1. INTRODUCTION

Quantum Monte Carlo (QMC) methods have shown promise in performance and scalings as an approach to quantum mechanical calculations.[1-5] These methods have been applied to the electronic structure, especially of solids[6-10] and medium-sized molecules,[11] but also to atoms and small molecules[12-28] presenting an alternative to traditional high-accuracy quantum chemistry methods such as configuration interaction (CI)[29] and coupled-cluster (CC).[30,31] Some typical limitations of standard *ab initio* methods are: 1) high-order dynamic electron correlation may be neglected or is prohibitively costly to compute, 2) scaling with the number of electrons is poor,[32,33] 3) some error may be introduced by internal contraction,[34] and 4) many algorithms are not yet efficient for large scale parallelization.

Key advantages for QMC methods are favorable scaling with the number of electrons and efficient parallelization (scaling with the number of computing cores). CI and CC methods can scale as n^7 (for n electrons) or worse[32] and thus rapidly become prohibitively expensive with increasing system size, whereas QMC methods scale as n^3 making them especially favorable for larger systems. In practice for small systems, despite the impressive n^3 scaling, QMC tends to have a large cost-prefactor making it relatively expensive compared with traditional quantum methods. Nevertheless, the vastly better scaling means that there exists a crossover point in system size beyond which QMC is favored. In addition, QMC algorithms can be very efficiently parallelized, scaling essentially linearly with the number of cores.[35] QMC methods can take full advantage of massively parallel machines, and are thus well-suited for next-generation computer architectures with millions of cores.[3,5]

During the development of QMC methods over the past few decades, there have been numerous reported studies of first-row atoms and (mostly) homonuclear diatomics[12-28,36] as well as hydrides.[37,38] The majority of those studies have focused on a single equilibrium geometry for each species while treating the total binding energy as well as components of the energy as method development benchmarks. It is well established that QMC methods can capture large fractions of both the strong and dynamic correlation energy, illustrated for example, by its impressive performance for the challenging Be_2 system.[17,25] Methods to compute forces have also been developed[39] and estimates of anharmonic force constants based on a few near-equilibrium points have been reported.[15]

Some PECs calculated with QMC have been reported[40-42] including most recently by Giner *et al.*[40] Those studies focus on technical aspects of the wavefunction construction and sampling as well as accuracy performance, but with less emphasis on assessments of cost and routine feasibility. They did achieve high-accuracy in comparison with reference curves, and reported that the fixed-node error (due to the fixed-node approximation)[43] was reduced as increasing numbers of determinants were used. Some of the studies didn't pursue the highest degree of accuracy, limiting the size of basis set in the trial wavefunction.[40] The largest bond dissociation distances were the least well converged relative to reference PECs,[37-40] and this was ascribed to limitations in the number of determinants.

For PECs, single-reference methods (with typical levels of truncation of the excitation operators) often do not produce correct behavior over the entire bond-distance range, because they cannot easily account for the evolution toward other configurations as the bond is stretched toward dissociation. The problem is common when breaking

multiple bonds, a well-known example being N_2.[44] Even though CCSD(T), *i.e.*, coupled-cluster with single, double, and perturbative triple excitations, is considered the "gold standard" of quantum chemistry, for N_2 it cannot accurately calculate the region of the PEC corresponding to dissociation. CCSD(T) breaks down at about twice the equilibrium bond length r_e and produces an artificial hump.[44, 45]

Multi-reference methods are often preferred as a globally applicable approach with correct dissociative behavior, and have been found necessary in previous QMC studies.[37] The single determinant FN-DMC atomization benchmarks reported by Grossman in 2002 found a 2.9 kcal/mol average absolute deviation over a 55 molecule test set, with deviations of 3.0 and 4.1 kcal/mol for CO and N_2 respectively.[14] The efficient use of multideterminants in QMC codes is not trivial but has been successfully addressed by an algorithm known as the *table method*[46,47] implemented in the QMCPACK[48] code package. The work reported here focuses on assessing the accuracy, cost, and routine feasibility of using QMC as implemented in two freely available packages, calculating ground state potential energy curves (PECs) for two test systems (CO and N_2). The multi-determinant trial wave functions used in this study were generated from orbitals and determinants using the multi-configurational self-consistent field (MCSCF) and configuration interaction (CI) methods from GAMESS (U.S.).[49] The CASINO[50] package was used for most of the QMC calculations reported here, along with some timing comparisons conducted using the QMCPACK[48] code.

2. FIXED-NODE QUANTUM MONTE CARLO

Here, we give a brief summary of some important aspects of QMC. For more technical details, the reader is referred to references 50 and 51.

2.1 VARIATIONAL MONTE CARLO

Variational Monte Carlo (VMC), using an approximate trial wave function Ψ_T, calculates the expectation value of a Hamiltonian *H*, with the integrals being performed by a Monte Carlo method.[52-54] The variational energy E_{VMC} is an upper bound to the exact ground state energy E_0 and is mathematically defined as

$$E_{VMC} = \frac{\int \Psi_T^*(\boldsymbol{R}) H \Psi_T(\boldsymbol{R})\, d\boldsymbol{R}}{\int \Psi_T^*(\boldsymbol{R}) \Psi_T(\boldsymbol{R})\, d\boldsymbol{R}} \geq E_0 \tag{1}$$

$\boldsymbol{R}$ is a 3*N*-dimensional vector of the coordinates ($\mathbf{r}_1$,$\mathbf{r}_2$,..,$\mathbf{r}_N$) of the *N* particles in the system (in this case electrons).[50] The expectation value of the Hamiltonian *H* can be rewritten with respect to the trial wave function Ψ_T as

$$\langle H \rangle = \frac{\int |\Psi_T(\boldsymbol{R})|^2 E_{loc}(\boldsymbol{R})\, d\boldsymbol{R}}{\int |\Psi_T(\boldsymbol{R})|^2\, d\boldsymbol{R}} \tag{2}$$

with the local energy $E_{loc} = \Psi_T^{-1} H \Psi_T$. By sampling the points $\boldsymbol{R}_i$ according to the distribution $|\Psi_T|^2$ with *M* configurations, the variational energy can then be computed from a set of local energies:

$$E_{VMC} = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} E_{loc}(\boldsymbol{R}_i) \tag{3}$$

Statistical uncertainty in a Monte Carlo method decreases as $1/\sqrt{M}$ where *M* is the number of samples. The primary purpose of the VMC method in this application is to optimize the parameters of a trial wave function, such as the Jastrow factor,[55-59] for subsequent use in the more accurate DMC (Diffusion Monte Carlo) method.

In this work, to help describe dynamic electron correlation, a three-body Jastrow factor was used,[52] which includes electron-electron *u* terms, electron-nucleus χ terms, and

electron-electron-nucleus f terms. The Jastrow factor makes the trial wave function depend explicitly on particle separations and is symmetric under the interchange of identical particles.[60] Note that in a similar application to the F_2 molecule, Giner *et al*. justified not employing a Jastrow factor, primarily in order to avoid the costly optimization of parameters.[40] They state that non-linear optimizations of Jastrow factors may make it more difficult to obtain smooth PECs. The main drawback to lacking a Jastrow factor conceded by Giner *et al*. is a greatly increased variance in the trial wavefunction and corresponding increased simulation times.[40] For a detailed description of the form of the Jastrow factor that is used in CASINO, see reference 52.

2.2 DIFFUSION MONTE CARLO

The Diffusion Monte Carlo (DMC) method[61,62] uses the importance-sampled imaginary-time Schrödinger equation to evolve an ensemble of electronic configurations toward the ground state. The Schrödinger equation can be written in integral form as

$$f(\boldsymbol{R},t) = \int G(\boldsymbol{R} \leftarrow \boldsymbol{R}', t-t') f(\boldsymbol{R}', t') d\boldsymbol{R}' \tag{4}$$

where $\boldsymbol{R}$ is a point in the configuration space of an N-particle system. Here, $f(\boldsymbol{R},t)$ is a mixed distribution dependent on some trial wave function Ψ_T, and written mathematically as $f(\boldsymbol{R},t) = \Psi(\boldsymbol{R},t)\Psi_T(\boldsymbol{R})$. The Green's function $G(\boldsymbol{R} \leftarrow \boldsymbol{R}', t-t')$ gives the probability of a transition from $\boldsymbol{R}$ to $\boldsymbol{R}'$ in the time interval $t-t'$, and satisfies the initial condition $G(\boldsymbol{R} \leftarrow \boldsymbol{R}', 0) = \delta(\boldsymbol{R} - \boldsymbol{R}')$.

Due to the fermion sign problem,[51] the DMC method in CASINO adopts the fixed-node approximation, in which the nodes of f are constrained to be the same as those of the trial wave function Ψ_T. The DMC method then produces the lowest energy possible for this nodal surface and can be considered variational with respect to it.[41]

Interestingly, for both atoms and diatomic molecules, Giner *et al*. reported a systematic decrease in the fixed-node error with respect to both the number of determinants retained and the size of the one-particle basis.[40]

2.3 COMPUTATIONAL METHOD

For all of the QMC calculations, multi-determinant Slater-Jastrow (MD-SJ) trial wave functions were used, which can be written mathematically as

$$\Psi_T(\boldsymbol{R}) = e^{J(\boldsymbol{R})} \sum_{k=1}^{N_D} c_k \, D_k^{\uparrow} D_k^{\downarrow} \tag{5}$$

where Ψ_T is the trial wave function, e^J is the Jastrow factor, c_k are the determinant coefficients for the multi-determinant expansion describing static correlation,[52] and $D_k^{\uparrow}$ and $D_k^{\downarrow}$ are the spin-up and spin-down Slater determinants, respectively. For both CO and N_2, the aug-cc-pwCV5Z basis sets[63] were truncated to 11s10p8d, *i.e.*, angular momentum functions $\geq f$ were removed. The orbitals in the trial wave functions were cusp corrected using the scheme of Ma *et al*.[21]

All DMC calculations were performed with the electrons moving one at a time (electron-by-electron). The time steps for all the systems were chosen so that the acceptance ratio of the proposed moves would be ~99.5%.

2.4 APPLICATIONS

For both systems, the Jastrow factor was defined using an expansion of order 8 for the u terms ($N_u = 8$), an expansion also of order 8 for the χ terms ($N_\chi = 8$), and an order of 4 for the f terms ($N_f^{en} = N_f^{ee} = 4$), resulting in a total of 149 variable parameters for N_2 and 281 parameters for CO in the two respective Jastrow factors. For all trial wave functions, the truncation order-parameter C, which determines the behavior at the

cutoff lengths, was set to a value of 3, providing a local energy that is continuous in configuration space.[60] For the electron-electron u terms and electron-electron-nucleus f terms, different parameter values were used for the parallel- and the antiparallel-spin electron-pairs, and, for the electron-nucleus χ terms, different parameters were used for spin-up and spin-down electrons.

The parameters of the Jastrow factor were optimized with unreweighted variance minimization. For each initial optimization cycle, the default cutoff lengths for the u, χ, and f terms were used and the initial Jastrow parameters were set to zero. To see how uncertainties in the optimized Jastrow factors would carry over to the subsequent DMC calculations, two procedures with vastly different costs were tested. *Procedure 1* is designed to seek high precision during the optimization stage, thus incurring a high computational cost. *Procedure 2* explores what compromises in accuracy are suffered when a less rigorous and time consuming optimization is performed. In *Procedure 1*, the Jastrow parameters were optimized using 5.0×10^5 configurations for one initial cycle, followed by a system-dependent number of additional optimization cycles. For N_2, the number of additional cycles was 6 and, for CO, the number of additional cycles was 9. In *Procedure 2*, the Jastrow parameters were optimized using 1.0×10^5 configurations (five times fewer than procedure 1) for 10 cycles with the cost of the optimizations being roughly an order of magnitude fewer CPU-hrs than *Procedure 1*. In both procedures, the subsequent DMC calculations were performed with target populations of 2500, a minimum of 1.5×10^5 sample points, and time steps of 0.002 a.u. for all geometries. For comparison with DMC, MRCI calculations were also done using the MOLPRO[64] package.

2.4.1 CO: $C(^3P_g) + O(^3P_g) \rightarrow CO(X^1\Sigma^+)$. When the ground states of carbon and oxygen atoms combine, the number of molecular states of CO, resolved into C_{2v} symmetry, is:[65]

$$^{5,3,1}(3\ A_1 + 2\ B_1 + 2\ B_2 + 2\ A_2)$$

which represents a total of nine states of each of three spin-multiplicities (singlet, triplet and quintet). Since there are nine singlet-states, dynamic weighting[66] was used to facilitate robust convergence near the asymptote for the full-valence (10*e*,8*o*), state-averaged multi-configurational self-consistent field (SA-MCSCF) calculations. As in the case of N_2, a high-spin (quintet) ROHF calculation was performed before the subsequent singlet DW-SA-MCSCF calculations.

To control the cost of the QMC calculations, only a limited number of determinants were retained (specified by two contending criteria). First, a coefficient cutoff of 0.002 was used at each point, such that all determinants with an absolute weight coefficient value > 0.002 were retained. This resulted in a varying number of retained determinants at each rCO distance. Secondly, for comparison, a fixed number of 250 determinants (those with the largest coefficients) were retained throughout the coordinate range.

To assess the quality of the QMC results for CO in comparison to typical high-level conventional electronic structure methods, an accurate MRCI-based reference PEC by Dawes *et al.*[67] was used. It was constructed from Davidson-corrected MRCI data, based on a dynamically-weighted state-averaged CASSCF reference (DW-SA-CASSCF), with a full-valence active space. The MRCI(Q) data was extrapolated to the complete basis set (CBS) limit with the aug-cc-pwCV*n*Z (*n* = 3-5) bases and all electrons

correlated. The vibrational levels on the PEC are of spectroscopic accuracy. The MRCI-based PEC also includes small spin-orbit (SO) and scalar-relativistic (SR) corrections. The small SO and SR correction terms were removed from the MRCI-based PES for the comparisons with DMC presented here. This permits a more direct comparison since the DMC Hamiltonian does not include those terms.

2.4.2 N_2: $N(^4S_u) + N(^4S_u) \rightarrow N_2(X^1\Sigma_g^+)$. When two ground state $N(^4S_u)$ nitrogen atoms combine, the total number of molecular states of N_2, resolved into D_{2h} symmetry, is $^{7,5,3,1}A_g$ (one state of each of four spin-multiplicities).[65]

For the calculations in GAMESS, to ensure convergence to the ground state of N_2, a high-spin (in this case septet) restricted-open Hartree-Fock (ROHF) calculation was performed at the largest N-N distance, followed by a 1-state MCSCF calculation for the desired singlet ground state with the full-valence (10*e*,8*o*) active space. In the QMC calculations, based on the results for CO discussed above, the strategy of retaining a variable number of determinants based on a coefficient threshold was abandoned, and a fixed (generous) number of 396 determinants were retained. The QMC data obtained along the coordinate range was assessed by comparison with an empirical spectroscopic PEC by Le Roy *et al*.[68]

3. RESULTS AND DISCUSSION

For CO, initially a coefficient weight cutoff of 0.002 was used to restrict the number of determinants retained in the trial wavefunctions, generally resulting in a different number of determinants employed at each CO bond distance. The DMC energy data and numbers of determinants are given in Tables 1 and 2 for *Procedure 1* and *Procedure 2* (respectively), which differ by roughly an order of magnitude in the

computational expense devoted to optimizing the Jastrow factor (the parameters are better converged by *Procedure 1*). The DMC energy data was overlaid with the reference PEC using weighted least-squares to account for the different uncertainties at each data point. The RMSD of the DMC energy data with respect to the reference PEC is 315 cm^{-1}. As shown in Figure 1, the points obtained via both (Jastrow optimization) procedures follow the reference PEC quite well near the equilibrium geometry. At bond distances > 2.0 Å where fewer determinants are retained, the DMC points tend to deviate slightly from the reference PEC (to higher energies). To see if this behavior might be related to the reduced numbers of determinants, trial wave functions with a fixed number of determinants (250) were optimized via both procedures and then recalculated with DMC. The results are given in Tables III and IV and in Figure 2. The RMSD of the DMC energy data computed using a fixed number of determinants with respect to the reference PEC is 292 cm^{-1}. This is only slightly less than that of the variable determinants data. Note that corresponding to the number of DMC samples, the uncertainties at each data point are in the range of 40-90 cm^{-1} (see Tables). It is not clear from these results that retaining a fixed number of determinants produces significantly more consistent results than the lower-cost coefficient weight cutoff strategy (which results in a variable number of retained determinants). Regarding the Jastrow optimization, since the results obtained by Procedures 1 and 2 are of similar quality and are essentially interchangeable, it appears that the investment of considerably more CPU time in *Procedure 1*, was not warranted.

To test whether sensible spectroscopic parameters could be obtained from the data, a fit to a Morse function was performed (it is straightforward to convert the Morse

parameters into anharmonic parameters).[15] First, since the most accurate PEC will not be precisely Morse-like, parameters obtained by fitting the reference PEC were compared to the experimental CO parameters of: r_e = 1.128 Å, ω_e = 2170.2 cm^{-1}, $\omega_e x_e$ = 13.46 cm^{-1}.[65] The Morse form could not accurately accommodate the entire coordinate range of the reference PEC and produced a fitted value for the harmonic constant that is significantly too high (ω_e = 2215.1 cm^{-1}). Fitting data in a more limited range of r_{CO} = [0.9, 1.6] Å, produced more reasonable values of r_e = 1.128 Å, ω_e = 2174.9 cm^{-1} and $\omega_e x_e$ = 13.05 cm^{-1}. Fitting the DMC data over the same coordinate range yields parameters of: r_e = 1.129 Å, ω_e = 2187.0 cm^{-1}, $\omega_e x_e$ = 13.28 cm^{-1}, all of which are quite useful estimates.

For N_2, a fixed number of 396 determinants was employed for all geometries. The data for Jastrow optimizations via *Procedure 1* and *Procedure 2* are given in Tables V and VI respectively, and are plotted in Figure 3. The data are compared with an empirical spectroscopic PEC by Le Roy *et al*.[68] The DMC data are generally consistent with the empirical PEC and are much closer to it than moderately high-level MRCI/AVTZ data shown for additional comparison. The RMSD of the DMC data with respect to the empirical PEC is 598 cm^{-1}, which is significantly larger than was found for CO discussed above. Since the empirical PEC for N_2 was obtained via direct fit to spectroscopic data, it implicitly includes small effects such as relativistic corrections not included in the DMC Hamiltonian (that were removed for this reason from the CO PEC for the previous comparison discussed above). However, this is not likely to be a significant source of discrepancy as these sort of corrections are relatively small for N_2. It is noteworthy that the DMC data point at r_{NN} = 0.8 Å (listed in Tables V and VI) was excluded from the comparison as it is an outlier roughly 5000 cm^{-1} more stable than the

value given by the reference PEC. The empirical PEC is expected to be unreliable for repulsive geometries far beyond the turning points of the contributing spectroscopic data. Indeed, the empirical PEC was confirmed to be much more repulsive at short range than a high quality *ab initio* PEC by Gdanitz which is consistent with the DMC value.[69] Again, as was found for CO, the additional cost of a stricter Jastrow optimization procedure did not yield obviously improved results.

As was done for CO, the reference and DMC-based PECs were each fit to a Morse function to extract spectroscopic parameters. The experimental values used for comparison are: r_e = 1.094 Å, ω_e = 2359.6 cm^{-1}, $\omega_e x_e$ = 14.46 cm^{-1}.[65] Fitting the reference PEC over the range r_{NN} = [0.9, 1.3] Å produced values of r_e = 1.098 Å, ω_e = 2360.6 cm^{-1}, $\omega_e x_e$ = 14.70 cm^{-1}. Fitting the DMC data over the same coordinate range yields parameters of: r_e = 1.097 Å, ω_e = 2373.4 cm^{-1}, $\omega_e x_e$ = 14.45 cm^{-1}, all of which are again quite accurate.

The costs of generating high quality energies via DMC for the two 14-electron systems (CO and N_2) were assessed. A significant fraction of the total cost relates to the optimization of a Jastrow factor that was employed in this study. As mentioned previously, some advantages and disadvantages of using a Jastrow have been noted by Giner *et al*.[40] The initial Jastrow optimizations for N_2 with a fixed number of determinants using cheaper *Procedure 2* took ~800-1300 CPU-hrs per point. With CO, for Jastrow optimizations via *Procedure 2* with a varying number of determinants (using a coefficient cutoff strategy), the CPU-hrs depend significantly on the number of retained determinants. The lowest cost was the bond distance of 3.5 Å, which includes 87 determinants and took about 300 CPU-hrs. The largest cost was 1.5 Å, which includes

317 determinants and took about 1860 CPU-hrs. For *Procedure 1*, the cost for CO increased to ~5,000-15,000 CPU-hrs per point, and the cost for N_2 increased to ~18,000-25,000 CPU-hrs per point.

Once the Jastrow has been optimized by VMC via either *Procedure 1* or *Procedure 2*, the subsequent DMC energy calculations add a significant additional cost which depends on the number of DMC samples which in turn determines the final uncertainties (related as $1/\sqrt{M}$ where M is the number of samples). For N_2, to reach uncertainties on the order of 100 cm^{-1} (see Tables), the DMC cost was ~1800-2200 CPU-hrs per point. Specifically, after Jastrow optimization via *Procedure 1*, the cost of the DMC sampling for the N_2 points located at r_{NN} = 0.8 Å, 1.1 Å, and 2.5 Å took 2170 CPU-hrs, 1850 CPU-hrs, and 2010 CPU-hrs, respectively, to reach an average uncertainty in the three points of ~86 cm^{-1}. Similarly, after *Procedure 2* optimization, the DMC cost for the same points was about 1880 CPU-hrs, 2035 CPU-hrs, and 2030 CPU-hrs, respectively, for an average uncertainty of ~93 cm^{-1}.

For CO, the DMC sampling cost after Jastrow optimization was ~775-1900 CPU-hrs per point to reach similar uncertainties. The cost of the DMC calculations for CO and N_2 (both 14-electron systems) was approximately the same, but the Jastrow optimizations via *Procedure 1* took about an order of magnitude longer than those of *Procedure 2*. The quality of the final energies are similar for the two procedures, indicating that a more conservative Jastrow optimization is reasonable since it appears that diminishing returns are realized with respect to further optimization. Trial wave functions for CO with 250 (fixed) determinants were also optimized using *Procedure 2* with a cost of ~1,000-1500

CPU-hrs per point. To achieve an average uncertainty ~33 cm^{-1} in the DMC, an additional cost of ~7,000-10,000 CPU-hrs per point was required.

The QMCPACK code was used to test the improved efficiency that is expected through use of the *table method* algorithm for multideterminant calculations.[46,47] The QMCPACK code was found to reach similar uncertainties for our two test systems in about 30% fewer CPU-hrs. For much larger numbers of determinants (up to 16,000), Clark *et al*. reported much more significant speedups in the range of factors of 15-40.[47] They noted smaller speedup factors for smaller numbers of determinants. The speedup that one might expect from the QMCPACK algorithm will depend on the size of the system and the number of determinants as well as hardware limitations such as memory and cache.[70] In our study of two small systems with only 14 electrons and modest numbers of determinants, the speedup is already significant indicating that this method should be preferred in future larger scale multideterminant applications.

Table 1. DMC energies for CO following *Procedure 1* (see text) for Jastrow factor optimization (data plotted in Figure 1).

Bond Distance (Å)	Number of determinants	DMC (a.u.)	uncertainty (+/-) (a.u)
0.90	107	-113.0951174	0.000413133
1.00	133	-113.2489414	0.000380672
1.10	153	-113.2959859	0.000415229
1.30	229	-113.2527814	0.000348658
2.60	204	-112.9007928	0.000193650
3.00	113	-112.8882332	0.000321403
3.30	95	-112.8846690	0.000296271
3.50	87	-112.8845678	0.000310887

Table 2. DMC energies for CO following *Procedure 2* (see text) for Jastrow factor optimization (data plotted in Figure 1).

Bond Distance (Å)	Number of determinants	DMC (a.u.)	uncertainty (+/-) (a.u.)
0.90	107	-113.0944380	0.000389325

Table 2. DMC energies for CO following *Procedure 2* (see text) for Jastrow factor optimization (data plotted in Figure 1). (cont.)

0.95	127	-113.1896378	0.000436312
1.00	133	-113.2494457	0.000436334
1.10	153	-113.2966373	0.000396878
1.20	202	-113.2880833	0.000415661
1.30	229	-113.2530581	0.000446796
1.50	317	-113.1619425	0.000352915
2.10	276	-112.9596469	0.000301529
2.50	229	-112.9078498	0.000321865
3.00	113	-112.8887365	0.000296142
3.30	95	-112.8856601	0.000307478
3.50	87	-112.8848458	0.000329048

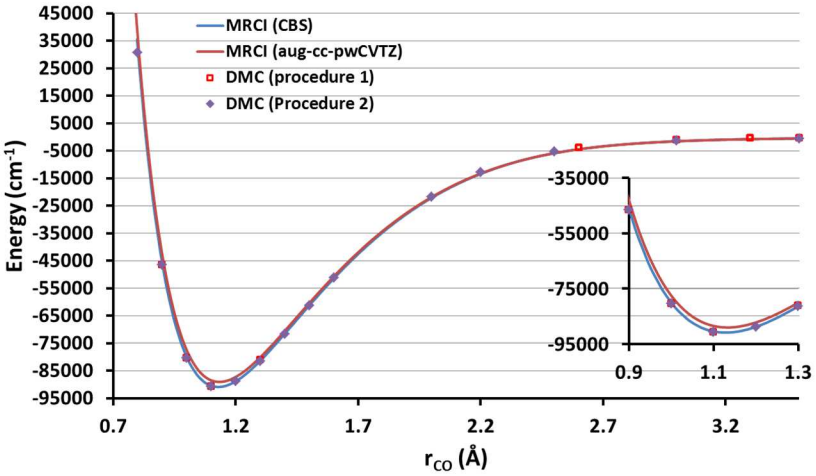

Figure 1. DMC calculations for CO with variable numbers of determinants are compared with two MRCI reference curves (see text).

Table 3. DMC energies for CO following *Procedure 1* Jastrow optimization with a fixed number of determinants.

Bond Distance (Å)	Number of determinants	DMC (a.u.)	uncertainty (+/-) (a.u.)
1.10	250	-113.2962092	0.000430335
1.20	250	-113.2874847	0.000453569
1.60	250	-113.1157087	0.000396781
2.20	250	-112.9419681	0.000350373
3.50	250	-112.8841369	0.000288665

Table 4. DMC energies for CO following *Procedure 2* Jastrow optimization with a fixed number of determinants.

Bond Distance (Å)	Number of determinants	DMC (a.u.)	uncertainty (+/-) (a.u.)
0.80	250	-112.7426924	0.000170137
0.90	250	-113.0946215	0.000182671
1.00	250	-113.2489800	0.000158881
1.10	250	-113.2961455	0.000167728
1.20	250	-113.2874901	0.000149456
1.30	250	-113.2537880	0.000168859
1.40	250	-113.2090661	0.000158616
1.50	250	-113.1623059	0.000164298
1.60	250	-113.1160313	0.000144766
2.00	250	-112.9814970	0.000139589
2.20	250	-112.9413138	0.000137728
2.50	250	-112.9071136	0.000138389
3.00	250	-112.8887795	0.000121424
3.50	250	-112.8852231	0.000126156

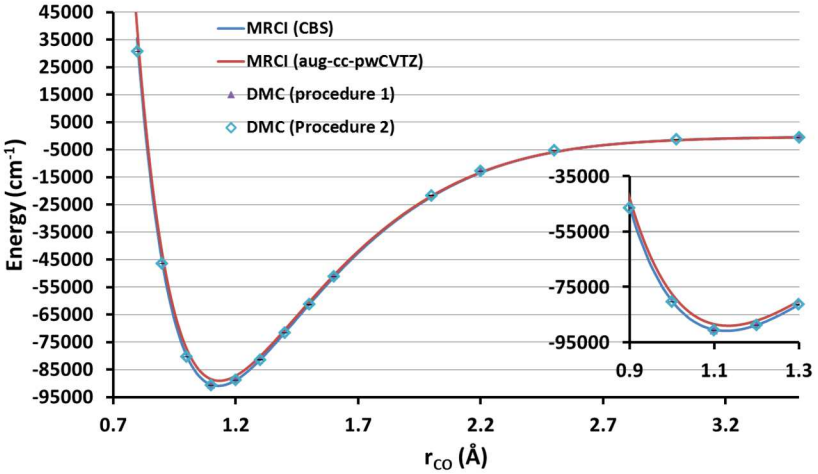

Figure 2. DMC calculations for CO with a fixed number (250) of determinants are compared with two MRCI reference curves (see text).

Table 5. DMC energies for N_2 following *Procedure 1* (see text) for Jastrow factor optimization (data plotted in Figure 3).

Bond Distance (Å)	Number of determinants	DMC energy (a.u.)	uncertainty (+/-) a.u.
0.80	396	-109.0167192	0.000434940
1.10	396	-109.5155949	0.000434508
1.20	396	-109.4944091	0.000447115
1.60	396	-109.2936484	0.000367967
1.80	396	-109.2220315	0.000335620

Table 5. DMC energies for N_2 following *Procedure 1* (see text) for Jastrow factor optimization (data plotted in Figure 3). (cont.)

2.00	396	-109.1811613	0.000554249
2.50	396	-109.1543457	0.000308338

Table 6. DMC energies for N_2 following *Procedure 2* (see text) for Jastrow factor optimization (data plotted in Figure 3).

Bond Distance (Å)	Number of determinants	DMC (a.u.)	uncertainty (+/-) (a.u.)
0.80	396	-109.0162778	0.000501235
0.90	396	-109.3448002	0.000453625
0.95	396	-109.4325595	0.000506966
1.10	396	-109.5155327	0.000455739
1.20	396	-109.4937945	0.000427817
1.30	396	-109.4478390	0.000430446
1.40	396	-109.3937795	0.000361739
1.50	396	-109.3405374	0.000365678
1.60	396	-109.2934221	0.000408069
2.00	396	-109.1806242	0.000341908
2.20	396	-109.1634908	0.000323882
2.50	396	-109.1548083	0.000323347

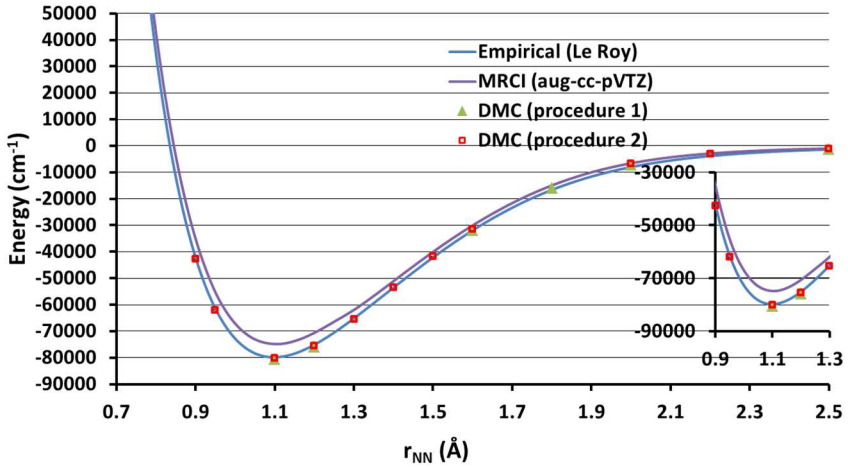

Figure 3. DMC calculations for N_2 compared with an MRCI curve and an empirical curve from Le Roy *et al*.[68]

4. CONCLUSION

It was determined that straightforward application of QMC methods implemented in two freely available codes (CASINO and QMCPACK) could robustly compute

electronic energies along dissociation coordinates of small molecules that are comparable in accuracy to high level traditional quantum chemistry. A QMC tutorial aimed at graduate students who have some familiarity with traditional quantum chemistry, but no experience with QMC is provided as *Supporting Information*.

Points along the potential energy curves of the ground states of CO and N_2 were generated with multideterminant fixed-node diffusion Monte Carlo methods and were found to be in close agreement with spectroscopically accurate curves. The spectroscopic constants obtained by fitting the data are in close agreement with experiment.

For the two 14-electron test systems, generating comparably high-quality electronic structure data by conventional methods such as MRCI takes less than 0.5 CPU-hr per point, compared with at least 4000 CPU-hr for the employed DMC method, as implemented in CASINO, depending on the desired final uncertainty. The QMCPACK code is known to be more efficient for multideterminant trial wavefunctions. For the small test systems and modest numbers of determinants employed in this study, the QMCPACK code was only slightly faster (~30%), but is indicated for larger scale applications where more significant speedups have been reported.[47] The favorable n^3 scaling of DMC does ensure a cross-over point in system size beyond which it becomes cheaper than conventional high-accuracy electronic structure methods (scaling as n^7 or worse). The large cost pre-factor of DMC seems to preclude it from routine use in the construction of global PESs (which for 3-5 atom systems typically require thousands of points) at this time. However, in addition to the favorable scaling with system size, QMC methods scale nearly linearly with the number of cores, which could lead to short time-to-solution using next generation architectures with millions of cores.

Even now, given the high accuracy that is achievable via QMC methods we also see it as a possible arbiter in difficult cases where high-level conventional methods might disagree about the presence or height of a rate determining reaction barrier.[71,72] It is anticipated that QMC methods will become increasingly relevant in the near future.

ACKNOWLEDGEMENTS

This work was supported by the U.S. Department of Energy [grant number DE-SC0010616]. RD thanks Paul Kent for useful discussions. AP thanks Albert DeFusco for help with interfacing GAMESS with CASINO.

REFERENCES

(1.) D. M. Ceperley and B. Alder, *Science* 231, 555 (1986).

(2.) D. M. Ceperley and L. Mitas. "Quantum Monte Carlo methods in chemistry." *New Methods in Computational Quantum Mechanics//Advances in Chemical Physics, XCIII, eds. Prigogine I. and Rice SA,(John Willcy & Sons, Inc. 1996)* 1-38.

(3.) K. P. Esler, J. Kim, D. M. Ceperley, W. Purwanto, E. J. Walter, H. Krakauer, S. Zhang, P. R. Kent, R. G. Hennig, C. Umrigar, and M. Bajdich, *J. Phys.: Conference Series,* 125, 012057 (2008).

(4.) B. M. Austin, D. Y. Zubarev, and W. A. Lester Jr. *Chem. Rev.* 112, 263 (2011).

(5.) A. Scemama, M. Caffarel, E. Oseret, and W. Jalby. " *J. Comp. Chem.* 34, 938 (2013).

(6.) A. Lüchow and J. B. Anderson. *Ann. Rev. Phys. Chem.* 51, 501 (2000).

(7.) J. Ma, D. Alfe, A. Michaelides, and E. Wang, *Phys. Rev. B* 80, 033407 (2009).

(8.) B. Santra, J. Klimes, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car and M. Scheffler, *Phys. Rev. Lett.* 107, 185701 (2011).

(9.) K. P. Esler, J. Kim, D. M. Ceperley, and L. Shulenburger, *Comp. in Sci. & Eng.* 14, 40 (2012).

(10.) N. Shannon, O. Sikora, F. Pollmann, K. Penc, and P. Fulde. *Phys. Rev. Lett.* 108, 067204 (2012).

(11.) I. G. Gurtubay, N. D. Drummond, M. D. Towler and R. J. Needs, *J. Chem. Phys.* 124, 024318 (2006).

(12.) C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. *Phys. Rev. Lett.* 60, 1719 (1988).

(13.) Z. Sun, R. N. Barnett and W. A. Lester Jr. *J. Chem. Phys.* 96*,* 2422 (1992).

(14.) J. C. Grossman, *J. Chem. Phys.* 117, 1434 (2002).

(15.) M. W. Lee, M. Mella, and A. M. Rappe, *J. Chem. Phys.* 122, 244103 (2005).

(16.) A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. B* 71, 066704 (2005).

(17.) J. A. W. Harkless and K. K. Irikura. *Int. J. Quant. Chem.* 106, 2373 (2006).

(18.) M. D. Brown, J. R. Trail, P. López Ríos and R. J. Needs. *J. Chem. Phys.* 126, 224110 (2007).

(19.) W. Purwanto, W. A. Al-Saidi, H. Krakauer and S. Zhang. *J. Chem. Phys.* 128, 114309 (2008).

(20.) N. Nemec, M. D. Towler, and R. J. Needs, *J. Chem. Phys.* 132, 034111 (2010).

(21.) B. Braïda, J. Toulouse, M. Caffarel and C. J. Umrigar. *J. Chem. Phys.* 134, 084108 (2011).

(22.) P. Seth, P. L. Ríos, and R. J. Needs, *J. Chem. Phys.* 134, 084105 (2011).

(23.) K. Hongo, and R. Maezono. *Int. J. Quan. Chem.* 112, 1243 (2012).

(24.) M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria. *J. Chem. Theo. Comp.* 8, 2181 (2012).

(25.) M. J. Deible, M. Kessler, K. E. Gasperich and K. D. Jordan, *J. Chem. Phys.*, 143, 084116 (2015).

(26.) R. C. Clay III and M. A. Morales. *J. Chem. Phys.*, 142, 234103 (2015).

(27.) Y. Yang, I. Kylänpää, N. M. Tubman, J. T. Krogel, S. Hammes-Schiffer and D. M. Ceperley. *J. Chem. Phys.*, 143, 124308 (2015).

(28.) R. Nazarov, L. Shulenburger, M. A. Morales and R. Q. Hood. *Phys. Rev. B* 93, 094111 (2016).

(29.) H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* 89, 5803 (1988).

(30.) R. J. Bartlett, *Ann. Rev. Phys. Chem.* 32, 359 (1981).

(31.) R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* 79, 291 (2007).

(32.) M. Musial and R. J. Bartlett, *J. Chem. Phys.* 122, 224102 (2005).

(33.) R. J. Bartlett and M. Musial, *Rev. Mod. Phys.* 79, 291 (2007).

(34.) L. Harding, S. Klippenstein, H. Lischka, and R. Shepard, *Theor. Chem. Acc.* 133, 1429 (2013).

(35.) M. D. Towler, *Psi-k Newsletter* 60, 166 (2003).

(36.) D. Cleland, G. H. Booth, C. Overy and A. Alavi. *J. Chem. Theo. Comp.* 8, 4138 (2012).

(37.) A. Lüchow and J. B. Anderson, *J. Chem. Phys.* 105, 4636 (1996).

(38.) S. Manten and Arne Lüchow, *J. Chem. Phys.* 115, 5362 (2001).

(39.) J. Vrbik and S. M. Rothstein, *J. Chem. Phys.* 96, 2071 (1992).

(40.) E. Giner, A. Schemama, and M. Carffarel, *J. Chem. Phys.* 142, 044115 (2015).

(41.) R. Springall, M. C. Per, S. P. Russo, and I. K. Snook, *J. Chem. Phys.* 128, 114308 (2008).

(42.) J. Toulouse and C. J. Umrigar, *J. Chem. Phys.* 128, 174101 (2008).

(43.) J. B. Anderson, *J. Chem. Phys.* 65, 4121 (1976).

(44.) X. Li and J. Paldus, *J. Chem. Phys.* 129, 054104 (2008).

(45.) W. D. Laidig, P. Saxe, and R. J. Bartlett, *J. Chem. Phys.* 86, 887 (1987).

(46.) P. KVV Nukala and P. R. C. Kent. *J. Chem. Phys.* 130, 204105 (2009).

(47.) B. K. Clark, M. A. Morales, J. McMinis, J. Kim and G. E. Scuseria, *J. Chem. Phys.* 135, 244105 (2011).

(48.) K. P. Esler, J. Kim, D. M. Ceperley, and L. Shulenburger,*Comput. Sci. Eng*. 14, 40 (2012).

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

(50.) R. J. Needs, M. D. Towler, N.D. Drummond and P. López Ríos, *J. Phys.: Condensed Matter* 22, 023201 (2010).

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

(52.) V. R. Pandharipande, S. C. Pieper and R. B. Wiringa, *Phys. Rev. B* 34, 4571 (1986).

(53.) S. Sorella, *Phys. Rev. B* 71, 241103 (2005).

(54.) N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. B* 70, 235119 (2004).

(55.) H.-J. Flad and A. Savin, *J. Chem. Phys.* 103, 691 (1995).

(56.) C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Phys. Rev. Lett.* 98, 110201 (2007).

(57.) E. Neuscamman, H. Changlani, J. Kinder and G. K.-L. Chan. *Phys. Rev. B* 84, 205132 (2011).

(58.) E. Neuscamman, C. J. Umrigar and G. K.-L. Chan. *Phys. Rev. B* 85, 045103 (2012).

(59.) E. Neuscamman, *J. Chem. Phys.* 139, 181101 (2013).

(60.) A. Ma, M. D. Towler, N. D. Drummond and R. J. Needs, *J. Chem. Phys.* 122, 224322 (2005).

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

(62.) C. J. Umrigar, M. P. Nightingale and K. J. Runge, *J. Chem. Phys.* 99, 2865 (1993).

(63.) K. A. Peterson and T. H. Dunning Jr., *J. Chem. Phys.* 117, 10548 (2002).

(64.) H.-J. Werner, P. J. Knowles, G. Knizia *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.

(65.) G. Herzberg, *Molecular Spectra and Molecular Structure. III. Electronic Spectra and Electronic Structure of Polyatomic Molecules* (D. Van Nostrand, 1966).

(66.) R. Dawes, A. W. Jasper, Chong Tao, C. Richmond, C. Mukarakate, S. H. Kable, and S. A. Reid, *J. Phys. Chem Lett.* 1, 641 (2010).

(67.) R. Dawes, X.-G. Wang, and T. Carrington, Jr., *J. Phys. Chem. A* 117, 7612 (2013).

(68.) R. J. Le Roy, Y. Huang, and C. Jary, *J. Chem. Phy.* 125, 164310 (2006).

(69.) R. J. Gdanitz, *Chem. Phys. Lett.* 283, 253 (1998).

(70.) Personal communication with P. R. C. Kent, 10/10/2016.

(71.) T. L. Nguyen, J. Li, R. Dawes, J. F. Stanton, and H. Guo, *J. Chem. Phys.* 117, 8864 (2013).

(72.) R. Dawes, P. Lolur, A. Li, B. Jiang, and H. Guo, *J. Chem. Phys.* 139, 201103 (2013).

III. INVESTIGATION OF THE OZONE FORMATION REACTION PATHWAY: COMPARISONS OF FULL CONFIGURATION INTERACTION QUANTUM MONTE CARLO AND FIXED-NODE DIFFUSION MONTE CARLO WITH CONTRACTED AND UNCONTRACTED MRCI

Andrew D. Powell,[1] Nikesh S. Dattani,[2] Rene F. K. Spada,[3] Francisco B. C. Machado,[3] Hans Lischka,[4] and Richard Dawes,[1,*]

[1] *Department of Chemistry, Missouri University of Science and Technology, Rolla, MO 65409*

[2] *Department of Chemistry, Kyoto University, Kyoto, Japan*

[3] *Departamento de Química, Instituto Tecnológico de Aeronáutica, São José dos Campos, 12.228-900 São Paulo, Brazil*

[4] *Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX 79409*

ABSTRACT

The association/dissociation reaction path for ozone ($O_2 + O \leftrightarrow O_3$) is notoriously difficult to describe accurately using ab initio electronic structure theory, due to the importance of both strong and dynamic electron correlation. Experimentally, spectroscopic studies of the highest lying recorded vibrational states combined with the observed negative temperature dependence of the kinetics of oxygen isotope exchange reactions confirm that the reaction is barrierless, consistent with the latest potential energy surfaces. Previously reported potentials based on Davidson-corrected internally contracted MRCI, suffer a spurious reef feature in the entrance channel even at the complete basis set limit. Here, we report an analysis of comparisons between a variety of electronic structure methods including internally contracted and uncontracted MRCI (with and without Davidson corrections), as well as full configuration interaction quantum Monte Carlo, fixed-node diffusion Monte Carlo and Density Matrix Renormalization Group.

keywords: Ozone, MRCI, Quantum Monte Carlo, internal contraction

1. INTRODUCTION

Ozone plays several crucial roles in the atmosphere,[1] including protecting life from harmful UV radiation, as well as participating in reactions with a number of trace gases.[2-4] Signatures of ozone's usual isotope dynamics are imparted into other species and can provide insight into modern hydrological cycles and the dynamics of stratosphere-troposphere exchange, and through biogeochemical pathways can also tell us about Earth's environment millions of years ago.[5]

The measured thermal rate coefficients of ozone isotope exchange reactions

$$^{x}O + {}^{y}O^{z}O \rightarrow O_3^{*} \rightarrow {}^{z}O + {}^{x}O^{y}O \text{ or } {}^{y}O + {}^{x}O^{z}O, \quad (1)$$

where *x*, *y*, and *z* represent the different ^{16}O, ^{17}O and ^{18}O masses respectively, have steep negative temperature dependencies indicative of a barrierless mechanism to form the O_3^* complex.[6-9] Strong kinetic isotope effects (KIE) have been observed in these nearly thermoneutral exchange processes, leading to mass-independent fractionation (MIF) in the stratosphere.[10-13] The dynamics are non-statistical and the relative roles of differences in zero-point energy, symmetry and nuclear spin-statistics, unbound resonances, competing stabilization processes, as well as nonadiabatic effects such as spin-orbit and derivative coupling, and geometric phase effects are still under investigation. Spectroscopic evidence also supports a barrierless topography of the potential energy surface (PES). Recent measurements of progressions of vibrational levels approaching the highest-lying bound states, combined with theoretical analysis by Tyuterev *et al*., are much more consistent with a barrierless PES.[14]

Efforts to construct a PES for ozone useful to investigate some of the above-mentioned dynamical processes are constrained and guided by the large number of

recorded vibrational levels for various isotopologues as well as by the dissociation energy which is specified to remarkable precision (D_0 = 8563 ± 3.5 cm^{-1}) by the Argonne active thermochemical tables approach (ATcT),[15] which infers D_e ~ 9219 ± 10 cm^{-1}. The long range interaction between well-separated O_2 and O-atom fragments has been characterized by Lepers *et al*.[16] The region that is least well-defined is the transition region just below dissociation. No vibrational levels have so far been recorded within 600 cm^{-1} of dissociation, yet the topography in the uppermost part of the wells crucially determines the dynamics and kinetics, including the observed negative temperature dependence of the exchange reactions.

An accurate global potential energy surface (PES) for the ground electronic state of ozone was published in by Dawes *et al*. in 2013.[17] The PES lacks the spurious reef feature found in several previous studies[18-20] and was used in wavepacket based quantum scattering calculations to successfully reproduce the negative temperature dependence of exchange rate coefficients as well as the large KIEs.[21-24] More recently the PES was used in studies of the total number of bound vibrational states, their symmetry, and their density as a function of energy.[25] Prior to constructing the PES, an understanding of the origin of the reef feature and its sensitivity to calculational parameters was sought.

Preliminary insight into the origin of the spurious reef was reported in a 2011 study that showed that the height and even the existence of the barrier depend on the details of the *ab initio* calculations.[26] The reef was attributed to a widely avoided crossing in the transition region between the lowest lying excited $^1A'$ state and the ground molecular state connecting to the $O_2(^3\Sigma_g^-)$ + $O(^3P)$ asymptote. If adiabatic dissociation of ozone is considered,

$$O_3(^1A_1) \rightarrow O_2(^3\Sigma_g^-) + O(^3P_g) \quad (2)$$

three-fold degeneracy is reached asymptotically, represented as (2 A′ + A″) in the C_s symmetry group. Diabatically, the ground state of ozone connects to excited states of both atomic and molecular oxygen:

$$O_3(^1A_1) \rightarrow O_2(^1\Delta_g) + O(^1D_g) \quad (3)$$

which combine asymptotically to give (5 A′ + 5 A″). These two lowest singlet blocks combine to give a total of 13 singlet states (7A′, 6A″). In that study, to facilitate the switch in state character and to represent the asymptotically degenerate states on an equal footing, the 13 lowest singlet states were included in state-averaged complete active space self-consistent (SA-CASSSCF) calculations with dynamic weighting (DW).[27] Subsequent internally contracted *ic*MRCI calculations were found to be sensitive to several factors: 1) active space, 2) basis set completeness, 3) Davidson correction, 4) and perhaps the internal contraction error. The height of the reef feature was found to be lower using a full-valence active space (18*e*, 12*o*) than for a reduced space (12*e*, 9*o*) in which the 2*s* orbitals are closed (held doubly occupied). The height of the reef is progressively lower for each basis set in the correlation consistent series aug-cc-pV*n*Z (*n* = 3-6) approaching the CBS limit. The Davidson correction yields a significantly more attractive PES with a lower barrier and dissociation energy in much better agreement with experiment, confirming the importance of high-order dynamic correlation. However, despite all of that, results at the *ic*MRCI(Q)/CBS level (using a full-valence active space), still exhibit a very slight reef about 10 cm^{-1} in height. A 2013 PES reported by Ayouz and Babikov[28] fit to data at the *ic*MRCI(Q)/CBS level, used a single CASSCF reference state

with the reduced (12*e*, 9*o*) active space, and has a reef feature with a height of more than 100 cm^{-1}. The last important factor appeared to be internal contraction error in the MRCI calculations. Internal contraction has since been shown to introduce small but kinetically significant errors in the bottleneck region of the PES for other systems.[29,30] In the 2011 study, it was found that by using two reference states in the *ic*MRCI calculation, the transition was made slightly more attractive, just enough to be monotonically attractive (barrierless). The number of employed reference states affects, but does not systematically control the internal contraction error. Uncontracted *uc*MRCI calculations are much more computationally expensive and so no explicit tests were performed at that time. Nevertheless it was concluded that by considering the four factors listed above, a more realistic barrierless PES could be obtained.

In 2013, Tyuterev *et. al.*[31] published a spectroscopic PES fit to an analytic form describing one of the three global minimum isomer wells. This PES used a full valence (18*e*, 12*o*) CASSCF reference to calculate one-state Davidson-corrected MRCI data. The PES combines data from the AV5V basis with other data extrapolated to CBS(5,6) limit. The authors reported a submerged reef in their data, but produced two versions of the PES, one of which included a *Dawes correction* to remove the reef. They found that the reef feature or its absence affected the highest-lying vibrational levels. The PES without a reef produced much better agreement with experimental level progressions.[32]

The 2013 PES by Dawes *et al.*[17] was constructed using the insights reported in 2011. To promote convergence with respect to basis set completeness, the newly available explicitly-correlated multireference configuration interaction (MRCI-F12)[33] method was used. Explicitly-correlated F12 methods have proven to greatly improve

convergence with respect to basis set, and have been shown to provide near CBS quality with relatively small basis sets.[34] Using the VTZ-F12 basis directly produced a tiny barrier (~3 cm^{-1}), while bases at or beyond VQZ-F12 yield barrierless results. For the PES, the VQZ-F12 basis was used without further extrapolation as this best matched the dissociation energy. The full-valence (18*e*, 12*o*) active space was used, and to promote orbital stability for some stretched geometries, 20 singlet states (rather than 13) were included in the DW procedure. Further tests were conducted with respect to the number of reference states used in the *ic*MRCI-F12 calculations (which affects the internal contraction error as mentioned above). Ultimately, 7 reference states were used, (corresponding to all of the $^1A'$ states from the two lowest blocks), making the PES still more attractive in the transition region. The PES was used in time-dependent wave packet scattering calculations of the thermal rate constants for the O + O_2 isotope exchange reactions, which agree well with experiment, including their negative temperature dependence.[21-24]

Here, to more systematically determine the effect of internal contraction on the MEP, large basis set uncontracted MRCI (*uc*MRCI) calculations were performed with the COLUMBUS[35] program and compared with internally contracted MRCI (*ic*MRCI) calculations performed with the MOLPRO[36] program. In an effort to assess the high-order correlation contribution and role of the Davidson correction, initiator full configuration interaction quantum Monte Carlo (*i*-FCIQMC) calculations were performed with the NECI[37] code, and density matrix renormalization group (DMRG) calculations were performed using the BLOCK code[38] interfaced to MOLPRO. Finally, fixed-node diffusion Monte Carlo (DMC) calculations were performed with the quantum Monte

Carlo (QMC) package QMCPACK[39] to test the routine feasibility of benchmarking challenging reaction paths using that method. Section II briefly describes the Monte Carlo based methods, while Section III provides results and discussion, followed by a conclusion in section IV.

2. MONTE CARLO METHODS

Here we give a brief description of the different Quantum Monte Carlo methods used within this work.

2.1 QUANTUM MONTE CARLO (QMC)

Variational Monte Carlo (VMC) uses an approximate trial wave function Ψ_T, an initial reference for QMC, to calculate the expectation value of the Hamiltonian, the integration of which is performed by a Monte Carlo method.[40] In this study, VMC is primarily used to optimize parameters of a trial wave function for subsequent use in the more accurate diffusion Monte Carlo (DMC) method. DMC uses the importance-sampled imaginary-time Schrödinger equation to evolve an ensemble of electronic configurations toward the ground state.[41] Due to the fermion sign problem, DMC adopts the fixed-node approximation.[42] Since both VMC and DMC are Monte Carlo methods, statistical uncertainty decreases as $1/\sqrt{N}$ where N is the number of samples.

For all of the QMC calculations, multi-determinant Slater-Jastrow (MD-SJ) trial wave functions were used, which can be written mathematically as

$$\Psi_T(\boldsymbol{R}) = e^{J(\boldsymbol{R})} \sum_{k=1}^{N_D} c_k \, D_k^{\uparrow} D_k^{\downarrow} \tag{4}$$

where Ψ_T is the trial wave function, e^J is the Jastrow factor, c_k are the determinant coefficients for the multi-determinant expansion describing static correlation, and $D_k^{\uparrow}$ and

$D_k^{\downarrow}$ are the spin-up and spin-down Slater determinants, respectively. The multi-configurational trial wave functions were generated with the multi-configurational self-consistent field (MCSCF) and configuration interaction (CI) methods from GAMESS.[43] The aug-cc-pVQZ (AVQZ) basis set was used with the full valence (18*e*,12*o*) active space. Each trial wave function was combined with a three-body Jastrow factor containing electron-electron, electron-nucleus, and electron-electron-nucleus terms. An expansion order of 10 was used for the electron-electron and electrons nucleus terms, and an expansion order of 3 was used for the electron-electron-nucleus terms, resulting in a total of 82 optimizable parameters. VMC calculations with energy minimization were used to simultaneously optimize the Jastrow factor parameters and the coefficients of the configuration state functions (CSFs). For the initial optimization, the default cutoffs lengths were used with the initial Jastrow parameters set to zero. The Jastrow parameters and CSFs were optimized simultaneously for 10 cycles. The scheme of Ma *et al*.[44] was used to correct for the electron-nuclear cusps. Selecting the number of CSFs used in the trial wave functions involves balancing cost with accuracy. In this study, to limit the QMC cost, each trial wave function employed a fixed number of 750 CSFs with the largest coefficients. The DMC calculations were performed with a target population of 2880, and a time step of 0.0005 a.u. for all geometries.

2.2 FULL CONFIGURATION INTERACTION QUANTUM MONTE CARLO (FCIQMC)

FCIQMC is a quantum Monte Carlo method[45] designed to converge to the full configuration-interaction (FCI) energy, *i.e.*, the exact solution to the Schrödinger equation for a given basis set. Thus is contrast to DMC described in the previous subsection, results are directly comparable to those obtained by standard electronic

structure methods. The method simulates the imaginary-time Schrödinger equation of the interacting Hamiltonian based on a stochastic population dynamics of an evolving set of walkers which live and propagate in Slater determinant space. The method is able to converge to the FCI energy of a system once a system-dependent number of walkers is reached. Its initiator extension (*i*-FCIQMC) is designed to accelerate convergence to the FCI energy by reducing the number of walkers required.[46] Both FCIQMC and *i*-FCIQMC methods have been used to compute FCI energies in several benchmark studies.[47-49]

For this study, the NECI calculations were performed with the aug-cc-pVDZ (AVDZ) basis set.

3. RESULTS AND DISCUSSION

For all of the calculations, the O-O (O_2) fragment bond distance was fixed at its asymptotic equilibrium value of 2.282 a.u. The bond angle varies negligibly along the minimum energy path (MEP) in valence coordinates and so was held constant at 116.8°. Calculations were carried out for the dissociating bond at a series of distances between 3.60 a.u. and 4.95 a.u which cover the transition region. An optional correction for spin-orbit (SO) coupling reduces the dissociation energy of the reference PES by ~80 cm^{-1} at the asymptote. Since the SO correction was not applied to the energies compared here, the uncorrected PES with D_e = 9355 cm^{-1} was used in all comparisons. The long range region of the reference PES is consistent with other PESs and also with electrostatic treatments,[17,16] so to focus on the transition region, the zero of energy in the comparisons presented here was set at 4.95 a.u. along the dissociation coordinate. (This is ~243 cm^{-1} below the asymptote on the non-SO-corrected reference PES). At each point *ic*MRCI energies were computed with Molpro and *uc*MRCI energies were computed using the

Columbus code. For both the *ic*MRCI and the *uc*MRCI methods, a one-state full valence (18*e*, 12*o*) CASSCF reference wave function was used followed by a one-state MRCI calculation with the standard relaxed Davidson correction (Q). In all cases precise agreement was obtained between the two codes for the CASSCF reference energy. Three basis sets: aug-cc-pVDZ (AVDZ), aug-cc-pVTZ (AVTZ), and aug-cc-pVQZ (AVQZ) were used, and both the *ic*MRCI and the *uc*MRCI energies (with and without the Davidson correction) were extrapolated to the CBS limit (l^{-3} formula) using the AVTZ and AVQZ bases.

For *i*-FCIQMC, the restricted Hartree-Fock (RHF) method at the AVDZ basis level was used as the reference wave function. The simulation was initialized with a single walker on the reference determinant D_0, and then the calculation proceeded with a shift value of zero, allowing for an initial exponential growth of walkers. The initial time step was chosen to be 0.00014 a.u., which was allowed to be dynamically updated so that multiple walkers spawning from the same attempt would be rare. For the initiator method, a parameter n_a which governs which determinates to include in the initiator space, was selected as $n_a = 3$ (walkers). If $n_a = 0$, then all determinants would have been included in the initiator space. To reduce the memory requirements and CPU time due to a large number of additional spawned walkers, a cutoff κ of 0.01 was applied, meaning that the walkers with weights greater than κ would be left unchanged. For each point, the total number of walkers was initially grown to 8.0×10^6, after which point, a series of calculations using the semi-stochastic adaptation were performed, which contained 450,000 of the largest weighted determinants in the core space. The number of walkers was then doubled and subsequently followed by semi-stochastic calculations until the

total number of walkers reached 2.56 x 10^8. When the walkers reached this value, in order to reduce the CPU cost, the core space was reduced to 200,000 of the most dominant determinates. After a few semi-stochastic calculations, the number of walkers was then doubled again to 5.12x10^8 for all geometries. To determine how further increases in the total number of walkers could affect the FCIQMC energy, for the geometry of 3.75 a.u., the number of walkers was expanded to 1.0x10^9, then to 8.0x10^9 and even a final test at 16.0x10^9.

Table 1. Comparison of energies at AVDZ.

Geometry (Bohr)	*ic*MRCI	*ic*MRCI(Q)	*uc*MRCI	*uc*MRCI(Q)	*i*-FCIQMC
3.60	-224.90399	-224.94390	-224.91301	-224.95390	-224.947(1)
3.75	-224.90441	-224.94403	-224.91313	-224.95360	-224.9551(2)
4.05	-224.90554	-224.94467	-224.91373	-224.95345	-224.9489(3)
4.50	-224.90647	-224.94514	-224.91415	-224.95314	
4.95	-224.90629	-224.94466	-224.91373	-224.95227	-224.9520(2)

Table 2. Comparison of energies at AVTZ.

Geometry (Bohr)	*ic*MRCI	*ic*MRCI(Q)	*uc*MRCI	*uc*MRCI(Q)
3.60	-225.06075	-225.11240	-225.07141	-225.12394
3.75	-225.06063	-225.11191	-225.07096	-225.12298
4.05	-225.06122	-225.11189	-225.07097	-225.12208
4.50	-225.06201	-225.11213	-225.07123	-225.12148
4.95	-225.06191	-225.11166	-225.07068	-225.12034

Table 3. Comparison of energies at AVQZ.

Geometry (Bohr)	*ic*MRCI	*ic*MRCI(Q)	*uc*MRCI	*uc*MRCI(Q)	DMC
3.60	-225.11018	-225.16506	-225.12096	-225.17662	-225.3431(4)
3.75	-225.10997	-225.16445	-225.12042	-225.17553	-225.3437(4)
4.05	-225.11044	-225.16427	-225.12031	-225.17447	-225.341(1)
4.50	-225.11116	-225.16441	-225.12048	-225.17375	-225.343(1)
4.95	-225.11102	-225.16389	-225.11991	-225.17258	-225.3444(3)

Table 4. Comparison of energies at the CBS limit.

Geometry (Bohr)	*ic*MRCI	*ic*MRCI(Q)	*uc*MRCI	*uc*MRCI(Q)
3.60	-225.14625	-225.20348	-225.15712	-225.2150730
3.75	-225.14597	-225.20279	-225.15650	-225.2138893
4.05	-225.14636	-225.20250	-225.15631	-225.2127035
4.50	-225.14702	-225.20256	-225.15643	-225.2119044
4.95	-225.1468691	-225.2020169	-225.1558374	-225.2107143

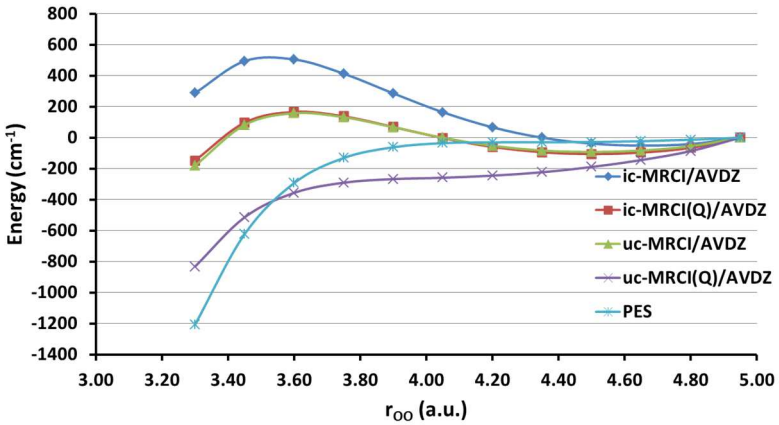

Figure 1. Comparison of *ic*MRCI, *uc*MRCI, *ic*MRCI(Q), and *uc*MRCI(Q) with the AVDZ basis set. Zero of energy is set at r_{OO} = 4.95 a.u., roughly 243 cm^{-1} below the asymptote (see text).

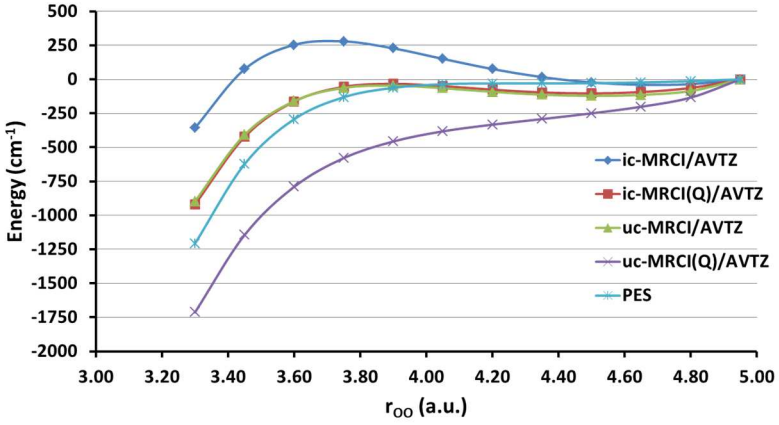

Figure 2. Comparison of *ic*MRCI, *uc*MRCI, *ic*MRCI(Q), and *uc*MRCI(Q) with the AVTZ basis set. Zero of energy is set at r_{OO} = 4.95 a.u., roughly 243 cm^{-1} below the asymptote (see text).

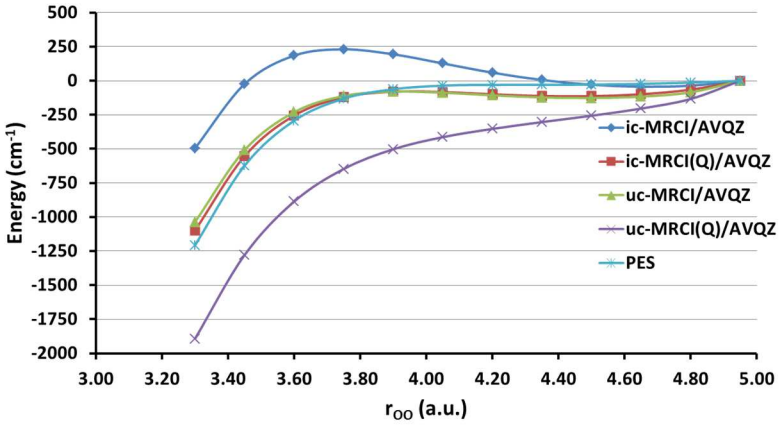

Figure 3. Comparison of *ic*MRCI, *uc*MRCI, *ic*MRCI(Q), and *uc*MRCI(Q) with the AVQZ basis set.

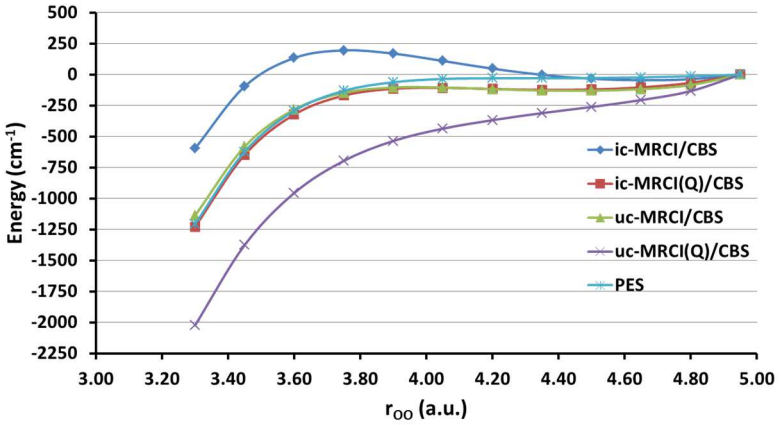

Figure 4. Comparison at the CBS limit.

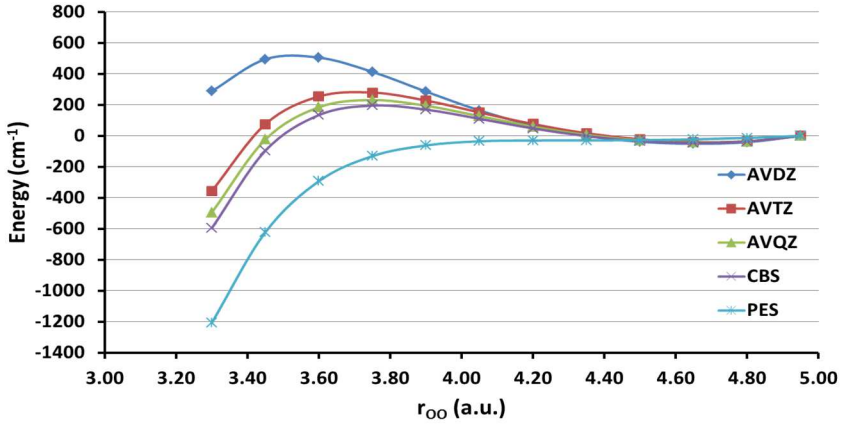

Figure 5. *ic*MRCI at different basis set levels.

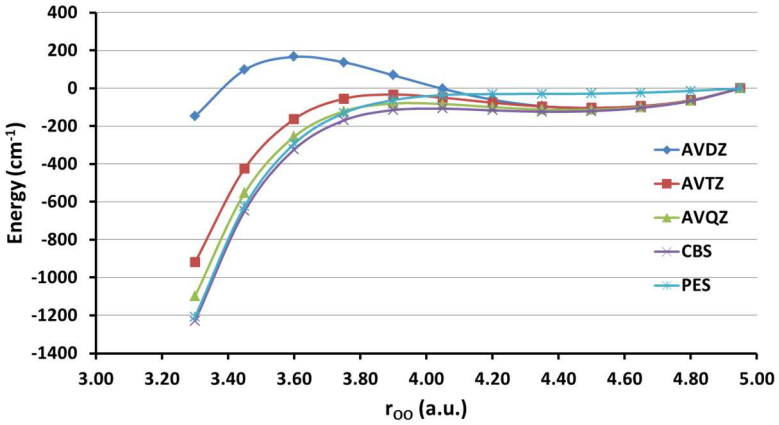

Figure 6. *ic*MRCI(Q) at different basis set levels.

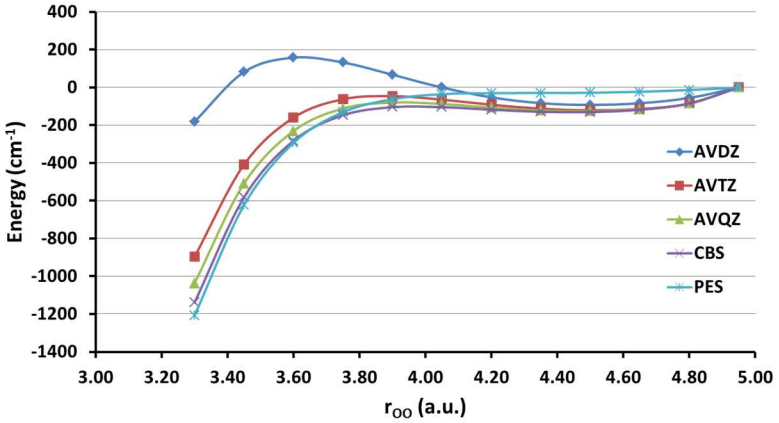

Figure 7. Comparison of ucMRCI at different basis set levels.

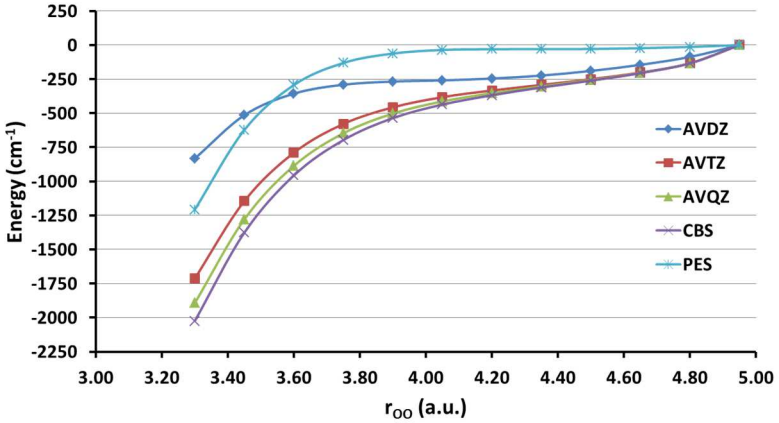

Figure 8. Comparison of ucMRCI(Q) at different basis set levels.

Table 5. *i*-FCIQMC results for the first nine most heavily weighted determinants. Total number of walkers = 2.56×10^8.

r_{OO} (a.u.)	Excitation level from reference determinant	Number of walkers on each determinant	Weight
3.75	0	14013	0.43040
	4	13073	0.40152
	1	7014	0.21543
	3	6521	0.20029
	2	5541	0.17021
	2	5158	0.15843
	1	3404	0.10456
	2	3244	0.09966
	4	3161	0.09709

Table 6. *i*-FCIQMC results for the first nine most heavily weighted determinants. Total number of walkers = 1.0×10^9.

r_{OO} (a.u.)	Excitation level from reference determinant	Number of walkers on each determinant	Weight
3.75	0	41579	0.45093
	4	38758	0.42034
	1	21357	0.23163
	3	20006	0.21697
	2	17841	0.19349
	2	17162	0.18613
	2	10992	0.11921
	2	10684	0.11588
	2	10626	0.11524

Table 7. *i*-FCIQMC results for the first nine most heavily weighted determinants. Total number of walkers = 5.12×10^8.

r_{OO} (a.u.)	Excitation level from reference determinant	Number of walkers on each determinant	Weight
3.60	0	25687	0.46118
	1	16140	0.28977
	4	15870	0.28493
	2	12540	0.22513
	3	10749	0.19299
	2	9769	0.17540
	2	8561	0.15371
	2	6811	0.12228
	1	6684	0.12000

Table 8. *i*-FCIQMC results for the first nine most heavily weighted determinants. Total number of walkers = 5.12×10^8.

r_{OO} (a.u.)	Excitation level from reference determinant	Number of walkers on each determinant	Weight
4.05	0	49497	0.71100
	1	15715	0.22574
	4	12114	0.17402
	1	10906	0.15667
	2	9624	0.13825
	1	8147	0.11704
	1	7493	0.10763
	1	6325	0.09085
	2	5759	0.08274

Table 9. *i*-FCIQMC results for the first nine most heavily weighted determinants. Total number of walkers = 5.12×10^8.

r_{OO} (a.u.)	Excitation level from reference Determinant	Number of walkers on each determinant	Weight
4.95	0	57077	0.64135
	4	44433	0.49928
	1	9930	0.11159
	1	9747	0.10953
	1	8935	0.10040
	4	7738	0.08696
	4	7559	0.08494
	1	7527	0.08459
	4	6938	0.07796

Table 10. *i*-FCIQMC results for the first 1,000 determinants with the excitation level and the total number of each excitation.

r_{OO} (a.u.)	Total number of walkers N_w	Excitation level from the reference determinant with number of such excitations					
		Single	Double	Triple	Quadruple	Quintuple	Sextuple
3.60	5.12×10^8	89	467	217	166	52	8
3.75	1.0×10^9	85	381	163	198	129	43
4.05	5.12×10^8	116	747	92	37	6	1
4.95	5.12×10^8	88	524	34	126	162	65

Table 11. Results for the first 10,000 determinants with the excitation level and the total number of each excitation.

r_{OO} (a.u.)	Excitation level from the reference determinant with number of such excitations					
	Single	Double	Triple	Quadruple	Quintuple	Sextuple

Table 11. Results for the first 10,000 determinants with the excitation level and the total number of each excitation. (cont.)

3.60	214	3578	1764	1813	1714	916

Energies computed with small basis, one-state *ic*MRCI/AVDZ exhibit a pronounced barrier with a height of ~556 cm^{-1}. [In this discussion we refer to the barrier height (if any) relative to what would be a vdW minimum were the barrier not spurious. Depending on the basis set and method, the barrier under discussion might be submerged with respect to the asymptote]. With the Davidson correction, the barrier height is reduced to ~272 cm^{-1}. Uncontracted *uc*MRCI/AVDZ without Davidson correction has a similar barrier as *ic*MRCI(Q) (with Davidson correction) with a barrier height of ~250 cm^{-1}. However, Davidson corrected *uc*MRCI(Q) is already barrierless even at the AVDZ basis level. For the AVDZ basis only, FCIQMC calculations with the initiator extension were used in an effort to benchmark the FCI limit and hence provide insight into the accuracy of the various calculations and the Davidson correction. Due their enormous cost (10s of thousands of CPU hours), all of the geometries except 3.75 a.u. were converged to 5.12×10^8 walkers with nominal statistical uncertainties on the order of 40 cm^{-1}. The *i*-FCIQMC energies at those points are between the *ic*MRCI(Q) and *uc*MRCI(Q) results, which might lead one to conclude that the Davidson correction slightly overcorrects in this case. However, it is noteworthy that the uncertainties given for the *i*-FCIQMC method in Table I should be interpreted cautiously. The derived uncertainties at intermediate stages of the calculations (smaller numbers of walkers) did not accurately reflect the range of where the energy might end up. In fact, in these results only the point at 3.75 a.u. appears to be truly well-converged. For that point, no further lowering of the energy was obtained beyond 8 billion walkers, but drops much larger

than the nominal uncertainty at earlier stages were noted upon each doubling of the walkers. For the point at 3.75 a.u. the *i*-FCIQMC energy is below that of all the other methods including the Davidson corrected uncontracted *uc*MRCI(Q) result. Thus, the best estimate of the FCI/AVDZ energy is below *uc*MRCI(Q), but unfortunately it was deemed too computationally expensive to converge all of the points to that same degree. The total CPU time to converge to 5.12×10^8 walkers for the points at 3.60, 4.05 and 4.50 a.u. was ~83,000 CPU-hrs, ~128,000 CPU-hrs, and ~103,000 CPU-hrs, respectively. In order to reach 8.0×10^9 walkers at 3.75 a.u., the computational cost was >175,000 CPU-hrs. Note that these results are for the AVDZ basis set and of course capturing a large fraction of the dynamic correlation energy requires much larger basis sets.

The *i*-FCIQMC method was able to provide insight into the multireference character of the electronic structure of ozone along the studied pathway. Note that the details of specific configurations important to the bonding description of ozone have been discussed before by Ruedenberg and coworkers.[50,51] Here we confine our discussion to the excitation levels and the implications for various computational approaches. At 3.75 a.u. with 2.56×10^8 walkers, the two most heavily weighted determinants were the HF reference and a configuration related by a quadruple excitation, both of which had nearly equal populations of ~14,000 walkers. As given in Table V, the next seven leading determinants, in terms of decreasing weights, were related by a single excitation, a triple excitation, two double excitations, another single excitation, a double excitation, and finally a quadruple excitation. At 1.0×10^9 walkers (Table VI), the first few leading configurations were the same with similar relative weights. At this number of walkers, considering the first 10,000 determinants, ~1800 determinants were found to be

quadruple excitations with respect to the HF reference, while ~1700 were quintuple excitations and ~900 were sextuple excitations. This highlights the challenge that ozone presents for single-reference based approaches. Indeed, it has been noted previously that the triples contribution shifts the harmonic frequencies by more than 100 cm^{-1}.[52] At the geometries of 3.60, 4.05, and 4.95 a.u. with $5.12x10^8$ walkers, many important contributions from quadruply excited configurations were also found. The second most heavily weighted determinant at 4.95 a.u. was a quadruple excitation, while at 3.60 and 4.05 a.u., it was found to be the third most weighted determinant. Interestingly at 4.95 a.u., of the top nine most heavily weighted determinants, four were quadruple excitations, and in the first 1,000 determinants, the number of quintuple excitations (~160) was found to be greater than the number of quadruple excitations (~130). These examples of important contributions from high excitation levels indicate why even triple-excitation single reference based methods will not suffice for ozone.

There has been some speculation about a significant strong/static correlation contribution from the 3*s* and 3*p* orbitals in this system. The cost to perform CASSCF calculations with larger than full-valence active spaces has been prohibitive for ozone even for small basis sets. Here to explore this issue, the DMRG method (as implemented in the Block program)[38] was used to perform CASSCF calculations also opening the 3*s* and 3*p* orbitals to construct an (18*e*, 24*o*) active space. The calculation was performed for a dissociating bond distance of 3.90 a.u. using the AVDZ basis set. The progression of energies is rather interesting. The HF/AVDZ energy of -224.2362 a.u. is lowered to -224.5089 a.u. for CASSCF with the usual full-valence active space (18*e*, 12*o*). This large drop is due to the fact that no particular configuration is very dominant and many

configurations contribute significantly. The leading squared coefficient is only 0.34, with the next few values being 0.22, 0.11, 0.10, 0.05 etc. However, the CASSCF energy for the (18*e*, 24*o*) active space including the 3*s* and 3*p* orbitals only lowers the energy to -224.5357, not as significant a drop as might be expected if the 3*s* and 3*p* orbitals contribute significantly to the bonding description. Note that the corresponding *ic*MRCI and *ic*MRCI(Q) values are -224.9050 and -224.9443 a.u. respectively (the *uc*MRCI and *uc*MRCI(Q) values are -224.9134 and -224.9535 a.u.). Overall this indicates that while strong correlation within the valence space is very important, strong correlation contributions from 3*s* and 3*p* orbitals are relatively insignificant. On the other hand, dynamic correlation is also very large and based on the large size of the Davidson correction, so too is the high-order contribution not directly captured in the MRCI(SD) treatment.

The *ic*MRCI has a significant barrier at each basis set with a barrier height of ~240 cm^{-1} remaining at CBS. Interestingly, the *ic*MRCI(Q) and *uc*MRCI show a similar difference in relative energies and have similar curves to each other at each basis level from AVDZ to the CBS limit (see Graphs I-IV). At CBS, *ic*MRCI(Q) still has a slight barrier of ~15 cm^{-1} while *uc*MRCI has a slightly larger barrier of ~25 cm^{-1}.

As the graphs show, the reef is strongly dependent upon the basis set. The *ic*MRCI(Q) has a barrier height of ~272 cm^{-1} for AVDZ, while at AVTZ, the is reduced to ~70 cm^{-1}. At CBS, the barrier nearly disappears completely (height of ~15 cm^{-1}). Similarly, *uc*MRCI has a barrier of ~250 cm^{-1} at AVDZ, but at CBS, the barrier is reduced to a height of ~25 cm^{-1}.

The Davidson corrected uncontracted *uc*MRCI(Q) does not produce a barrier at any basis set level. Thus, the use of internally contracted methods can certainly have a dynamically relevant impact. Similar to the reference PES, the *uc*MRCI(Q) has monotonically decreasing energies. However, the inclusion of the Davidson correction to the one-state *uc*MRCI(Q) calculation makes the method too attractive in this application when compared to the spectroscopically accurate PES. It will be of interest in the future to further characterize the PES at this level of theory, but with a balanced multistate treatment as was used to construct the PES. It is still prohibitively expensive to use *uc*MRCI(Q) (and large basis sets) to compute all of the points needed to construct a global PES.

The fixed-node DMC method seems to predict a barrier within the uncertainties, in contrast to the reference PES. However, in these tests only 750 CSFs were included in the trial wave functions, and it is not known at this time how sensitive the topography of the PES is to this restriction.

As an aside, both the restricted Hartree-Fock (RHF) and the configuration interaction with single and double excitations (CISD) methods, produce no submerged reef or barrier at any basis set level (in contrast to most of the MRCI methods).

4. CONCLUSION

Calculations were performed with internally contracted and uncontracted MRCI, *i*-FCIQMC, and fixed-node DMC along the association/dissociation MEP and were compared to a spectroscopically and dynamically accurate PES. Comparing *ic*MRCI and *uc*MRCI, it was found that internal contraction error indeed plays a significant role in producing the reef feature. One-state calculations with *ic*MRCI, *ic*MRCI(Q), and

*uc*MRCI all produce a barrier in contrast to the reference PES. The *uc*MRCI(Q) produced monotonically decreasing energies, but with respect to the PES, was too attractive, which would be inconsistent with the spectroscopy and dynamics. The *i*-FCIQMC method was used to benchmark energies at the AVDZ basis set. However, due to its very high computational cost, only the geometry at 3.75 a.u was fully converged. The best resulting energy at that point was found to be even lower than *uc*MRCI(Q), but the other points were not well enough converged to draw conclusions about the reef feature. The important configurations determined by the *i*-FCIQMC method reflect the multireference character of ozone, indicating important determinants of quadruple (and even sextuple) excitation levels from the single reference at geometries along the pathway. CASSCF calculations performed using the DMRG method with an active space expanded beyond full-valence to include the 3*s* and 3*p* orbitals, obtained a negligible strong correlation contribution from those orbitals. Thus, while strong correlation within the valence space is very important, strong correlation contributions from 3*s* and 3*p* orbitals are relatively insignificant. On the other hand, dynamic correlation is also very large and based on the large size of the Davidson correction, so too is the high-order contribution not directly captured in the MRCI(SD) treatment. These results highlight the challenging nature of ozone's electronic structure.

REFERENCES

(1.) R. P. Wayne, *Chemistry of Atmospheres* (Oxford, 1993).

(2.) C.A. Taatjes, *Annu. Rev. Phys. Chem.* 68, 183 (2017).

(3.) Y.-P. Lee, *J. Chem. Phys.* 143, 020901 (2015).

(4.) R. Dawes, B. Jiang, and H. Guo, *J. Am. Chem. Soc.*137, 50 (2014).

(5.) H. Bao, X. Cao, and J. A. Hayles, *Geochim. Cosmochim. Acta* 170, 39 (2015).

(6.) S. M. Anderson, F. S. Klein, and F. Kaufman, *J. Chem. Phys.* 83, 1648 (1985).

(7.) S. M. Anderson, D. Hulsebusch, and K. Mauersberger, *J. Chem. Phys.* 107, 5385 (1997).

(8.) M. R. Wiegell, N. W. Larsen, T. Pedersen, and H. Egsgaard, *Int. J. Chem. Kinet.* 29, 745 (1997).

(9.) P. Fleurat-Lessard, S. Y. Grebenshchikov, R. Schinke, C. Janssen, and D. Krankowsky, *J. Chem. Phys.* 119, 4700 (2003).

(10.) M. H. Thiemens, *Science* 293, 226 (2001).

(11.) M. H. Thiemens, *Annu. Rev. Earth Planet. Sci.* 34, 217 (2006).

(12.) K. Mauersberger, D. Krankowsky, C. Janssen, and R. Schinke, *Adv. At., Mol., Opt. Phys.* 50, 1 (2005).

(13.) Y. Q. Gao and R. A. Marcus, *Science* 293, 259 (2001).

(14.) V. G. Tyuterev, R. Kochanov, A. Campargue, S. Kassi, D. Mondelain, A. Barbe, E. Starikova, M. R. De Backer, P. G. Szalay, S. Tashkun, *Phys. Rev. Lett.* 113, 143002 (2014).

(15.) B. Ruscic and D. H. Bross, Active Thermochemical Tables (ATcT) values based on ver. 1.122 of the Thermochemical Network (2016); available at ATcT.anl.gov.

(16.) M. Lepers, B. Bussery-Honvault, and O. Dulieu, *J. Chem. Phys.* 137, 234305 (2012).

(17.) R. Dawes, P. Lolur, A. Li, B. Jian, and H. Guo, *J. Chem. Phys.* 139, 201103 (2013).

(18.) R. Hernández-Lamoneda, M. R. Salazar, and R. T. Pack, *Chem Phys. Lett.* 355, 478 (2002).

(19.) R. Schrinke and P. Fleurat-Lessard, *J. Chem. Phys.*, 121 5789 (2004).

(20.) F. Holka, P. G. Szalay, T. Müller, and V. G. Tyuterev, *J. Chem. Phys. A* 114, 9927 (2010).

(21.) Y. Li, Z. Sun, B. Jiang, D. Xie, R. Dawes, and H. Guo, *J. Chem. Phys.* 141, 081102 (2014).

(22.) Y. Li, Z. Sun, B. Jiang, D. Xie, R. Dawes, H. Guo, *J. Chem. Phys.* 141, 081102 (2014).

(23.) W. Xie, L. Liu, Z. Sun, H. Guo, R. Dawes, *J. Chem. Phys.* 142, 064308 (2015).

(24.) Z. Sun, D. Yu, W. Xie, J. Hou, R. Dawes, H. Guo, *J. Chem. Phys.* 142, 174312 (2015).

(25.) S. Ndengue, R. Dawes, X. Wang, T. Carrington Jr., Z. Sun and H. Guo, *J. Chem. Phys.* 144, 074302 (2016).

(26.) R. Dawes, P. Lolur, J.Mia, and H. Gao, *J. Chem. Phys.*, 135, 081102 (2011).

(27.) R. Dawes, A. W. Jasper, C. Tao, C. Richmond, C. Mukarakate, S. H. Kable, and S. A. Reid, *J. Phys. Chem. Lett.* 1, 641 (2010).

(28.) M. Ayouz and D. Babikov, *J. Chem. Phys.*, 138, 164311 (2013).

(29.) L. B. Harding, S. J. Klippenstein, H. Lischka, and R. Shepard, *Theor. Chem. Acc.* 133, 1429 (2014).

(30.) P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, *Chem. Rev.* 112, 108 (2012).

(31.) V. G. Tyuterev, R. V. Kochanov, S. A. Tashkun, F. Holka, and P. G. Szalay, *J. Chem. Phys.*, 139, 134307 (2013).

(32.) A. Barbe, S. Mikhailenko, E. Starikova, M.-R. DeBacker. V. G. Tyuterev, D. Mondelain, S. Kassi, A. Camparague, C. Janssen, S. Tashkun, R. Gamache, J. Orphal, *J. Quant. Spectrosc. Radiat. Transf.* 130, 172 (2013).

(33.) T. Shiozaki, G. Knizia, and H.-J. Werner, *J. Chem Phys.*, 134, 034113 (2011).

(34.) H.-J. Werner, T. B. Adler, G. Knizia, and F. R. Manby, *Recent Progress in Coupled-Cluster Methods*, edited by P. Cársky, J. Paldus, and J. Pittner (Springer, 2010).

(35.) H. Lischka, T. Müller, P. G. Szalay, R. M. Pitzer, and R. Shepard, *Comput Mol Sci* 1, 191 (2011).

(36.) H.-J. Werner, P. J. Knowles, G. Knizia *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.

(37.) The code can be obtained from https://github.com/ghb24/NECI_STABLE.

(38.) S. Sharma and G. K.-L. Chan. "Block code for dmrg." (2012).

(39.) J. Kim, K. P. Esler, J. McMinis, M. A. Morales, B. K. Clark, L. Shulenburger, and D. M. Ceperley, *J. Phys.: Conf. Ser.* 402, 012008 (2012).

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

(41.) R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, *J. Phys.: Condes. Matter* 22, 023201 (2010).

(42.) J. B. Anderson, *J. Chem. Phys.* 65, 4121 (1976).

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

(44.) A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs, *J. Chem. Phys*. 122, 224322 (2005).

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

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

(47.) G. H. Booth and A. Alavi, *J. Chem. Phys.*, 132, 174104 (2010).

(48.) G. H. Booth, D. Cleland, A. J. W Thom, and A. Alavi, *J. Chem. Phys.*, 135, 084104 (2011).

(49.) D. Cleland, G. H. Booth, C. Overy, and A. Alavi, *J. Chem. Theory Comput.*8, 4138 (2012).

(50.) V. A. Glezakou, S. T. Elbert, S. S. Xantheas, and K. Ruedenberg, *J. Physs Chem. A*, 114, 8923 (2010).

(51.) D. Theis, J. Ivanic, T. L. Windus, and K. Ruedenberg, *J. Chem. Phys.* 144, 104304 (2016).

(52.) J. D. Watts and R. J. Bartlett, *J. Chem. Phys*. 108, 2511 (1998).

SECTION

2. CONCLUSION

In this work, both traditional *ab initio* methods and quantum Monte Carlo methods were employed to solve the electronic Schrödinger equation for a variety of chemical systems. QMC was applied to potential energy surfaces, and in comparisons with *ab initio*, was generally found to be very accurate. The primary disadvantage of QMC is that it suffers from a large cost prefactor which make the calculations much more expensive than traditional methods for small systems. For comparable precision in the studied systems, traditional methods use <500 CPU-hrs, while QMC methods take thousands of CPU-hrs, as in the case of CO and N_2, or even up to hundreds of thousands of hours for ozone. Thus, QMC is not as applicable for routine use for small molecular systems as traditional methods. However, the algorithms of QMC are well parallelized, meaning that it can take advantage of thousands or hundreds of processors and reduce the time to solution, while most standard *ab initio* methods cannot. The time to solution could still be small depending on the number of processors used.

Since DMC has n^3 scaling with the number of electrons, it is often used for systems for which traditional high-accuracy *ab initio* methods cannot be used due to their expensive computational cost, *e.g.*, solids, since they scale as n^7 or higher (e. g., CCSD(T)). Also, FCIQMC can highlight some important aspects of electronic structure as in the case of ozone. These methods can take full advantage of modern architectures of computing resources containing millions of cores thus making the methods more relevant and applicable for future work.

REFERENCES

(1.) Reprinted with permission from L. Vereecken and J. S. Francisco, *Chem. Soc. Rev.*, 41, 6259 (2012). Copyright 2017.

(2.) J. C. Slater, *Rev. Mod. Phys*. 35, 484 (1963).

(3.) K. Raghavachari and J. B. Anderson, *J. Phys. Chem.*, 100, 12960 (1996).

(4.) J. A. Key and D. W. Ball, *Introductory Chemistry-1st Canadian Edition*. © 2014 Jessie A. Key and licensed under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/). From https://opentextbc.ca/. Accessed April 15, 2017.

(5.) http://www.kf.elf.stuba.sk/~ballo/piatok/prezentacia/hartree-fock/beyond_hf.html. Accessed April 12, 2017.

(6.) F. Jensen, *Introduction to Computational Chemistry*, (John Wiley & Sons, 2007).

(7.) K. L. Bak, P. Jørgenson, J. Olsen, T. Helgaker, and J. Gauss, *Chem. Phys. Lett.* 317, 116, (2000).

(8.) R. J. Bartlett and M. Musial, *Rev. Mod. Phys.* 79, 291 (2007).

(9.) K. Hilaro, *Recent Advances in Multireference Methods*, (World Scientific, 1999).

(10.) P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, *Chem. Rev.* 112, 108 (2012).

(11.) S. R. Langhoff and E. R. Davidson, *INt. J. Quantum Chem.*8, 61 (1974).

(12.) G. E. Moore, *Proc. SPIE* 2438, Advances in Resist Technology and Processing XII (1995).

(13.) Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported License (https://creativecommons.org/licenses/by-sa/3.0/deed.en). From https://en.wikipedia.org/wiki/Moore%27s_law. Accessed April 13, 2017.

(14.) C. P. Robert, *Monte Carlo Methods,* (Wiley StatsRef: Statistics Reference Online, 2016).

(15.) N. Makri and W. H. Miller, *Chem. Phys. Lett.*, 139, 10 (1987).

(16.) D. Ceperly and B. Alder, *Science*, 231, 555 (1986).

(17.) B. M. Austin, D. Y. Zubarev, and W. A. Lester, Jr. *Chem. Rev.* 112, 263 (2012).

(18.) D. Alfè, A. P. Bartók, G. Csányi, and M. J. Gillan, *J. Chem. Phys.*138, 221102 (2013).

(19.) J. A. Santan, J. T. Krogel, J. Kim, P. R. C. Kent, and F. A. Reboredo, *J. Chem. Phys.* 142, 164705 (2015).

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

(21.) N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. B* 70, 235119 (2004).

(22.) S. R. Chinnamsetty, H. Luo, W. Hackbusch, H.-J. Flad, A. Uschmajew, *Chem. Phys.,* 401, 36 (2012).

(23.) H. Luo, W. Hackbusch, and H.-J. Flad, *Mol. Phys.*, 108, 425 (2010).

(24.) K. Raghavachari and J. B. Anderson, *J. Phys. Chem.*, 100 12960 (1996).

VITA

Andrew D. Powell, in May of 2011, received two Bachelor of Science degrees, Chemistry and Mathematics, from the College of the Ozarks, Point Lookout, Missouri, USA. In July 2017, he received his Ph.D. in Chemistry from the Missouri University of Science and Technology, Rolla, Missouri, USA.

He has published journal articles and presented his work at various conferences, some of which are presented in this work. In March 2011, he presented his undergraduate research at the national meeting of the American Chemical Society (ACS). In June 2014, he presented at the 69th International Symposium for Molecular Spectroscopy (ISMS), and in June 2016, he also presented at the 71st International Symposium for Molecular Spectroscopy. In April 2017, he won the graduate student excellence award for the chemistry department.

In 2010, he was inducted as a member of Sigma Zeta, and he has been a member of the American Chemical Society since 2011.