Investigating the Performance of Perturbative Excitations in Coupled Cluster Theory

by

James Matthew Gabertan Begley

(Under the Direction of Henry F. Schaefer III)

Abstract

Ab initio quantum chemistry is an incredibly useful tool to compute highly accurate energies, frequencies, and molecular geometries for a plethora of chemical systems, but the great computational cost associated with such accurate methods hinders the efficacy of its use without cost-reductive techniques. Perhaps one of the most prominent *ab initio* methods, coupled cluster theory with perturbative excitations can approximate highly accurate energies on small molecular systems while greatly decreasing the computational cost compared to coupled cluster implementations of the same order. In this dissertation, the theory behind various *ab initio* quantum chemical methods, including coupled cluster theory and perturbation theory, is discussed. These methods are then employed to compute the enthalpies of formation for a set of Criegee intermediates, or carbonyl oxides. Additionally, coupled cluster theory is employed on second-row homonuclear diatomic molecules to compute their potential energy surfaces and relevant spectroscopic constants. The convergence of coupled cluster methods on each of these systems is explored along with the performance of perturbative excitation methods, including the popular CCSD(T) and CCSDT(Q) methods. Through these studies, it is shown that the use of perturbative excitations in coupled cluster theory should be used with caution on systems with significant multi-reference character. Otherwise,

these methods are proven to be accurate, cost-effective approximations to full coupled cluster theory computations of the same order.

INDEX WORDS: Electronic Structure Theory, Hartree-Fock, Coupled Cluster Theory, Perturbation Theory, Focal Point Approach, Electron Correlation, Criegee Intermediates, Diatomic Molecules

Investigating the Performance of Perturbative Excitations in Coupled Cluster Theory

by

James Matthew Gabertan Begley

B.S., Clemson University, 2020

A Dissertation Submitted to the Graduate Faculty of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2025

Investigating the Performance of Perturbative Excitations in Coupled Cluster Theory

by

James Matthew Gabertan Begley

Major Professor: Henry F. Schaefer III

Committee: Suzanne R. Ellenberger
Brandon Rotavera

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2025

Dedication

To my parents and my brother, for their never-ending support through my endeavors in graduate school. You are the reason I am still here today.

Acknowledgments

I would like to thank my advisor, Dr. Schaefer, for the unending support he has given me throughout my graduate career and for fostering a one-of-a-kind community in his research group. I must also thank Dr. Turney for his guiding hand in my research and keeping me grounded during my time here. I would also like to thank the rest of the members of the CCQC for the support, the laughs, the camaraderie, and (most importantly) the criticism.

Table of Contents

Chapter 1

Introduction and Theory

Quantum chemistry is a powerful tool that has been utilized by experimental chemists for decades, both to confirm observations from laboratory experiments as well as to predict future ones. As time progresses, the challenges that face quantum chemists are to obtain more accurate computations, faster, for larger systems. These are three attributes that are rarely obtained simultaneously; as one pursues more accurate computations, the computational cost of obtaining the data increases and the data cannot be feasibly computed on larger systems. The inverse is true as well: as one attempts to compute data faster on larger molecules, the accuracy of such computations will almost certainly decrease. So, the most pressing issue for a computational chemist is how to obtain the most accurate data possible on relevant chemical systems without increasing computational cost to intractable levels.

In Chapter 1, relevant quantum chemical methods are discussed. Unless otherwise stated, the information presented in this chapter comes from a set of textbook sources.[1,2] In subsequent chapters, the ability for approximations to high-order computational methods to describe accurate data is studied.

1.1 Schrödinger Equation

At its very core, the objective of computational quantum chemistry is to solve the time-dependent Schrödinger equation:

$$\hat{H}\Psi(x,t) = i\hbar\frac{\delta}{\delta t}\Psi(x,t) \tag{1.1}$$

in which $\hat{H}$ denotes the Hamiltonian operator and Ψ is the wavefunction. The Hamiltonian may be split into kinetic ($\hat{T}$) and potential ($\hat{V}$) components, such that

$$\hat{H} = \hat{T} + \hat{V} \tag{1.2}$$

Assuming that the potential is time-independent, the wavefunction may be rewritten as

$$\Psi(x,t) = \Psi(x)e^{-\frac{iEt}{\hbar}} \tag{1.3}$$

such that E represents the energy of the system. Therefore, it follows that the time-independent Schrödinger equation may be written, generally, as:

$$\hat{H}\Psi = E\Psi \tag{1.4}$$

The Hamiltonian operator of the time-independent Schrödinger equation may be split into multiple parts:

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_{eN} + \hat{V}_{NN} + \hat{V}_{ee} \tag{1.5}$$

These parts, respectively, describe the kinetic energy of the electrons, the kinetic energy of the nuclei, the electron-nucleus attraction, the nucleus-nucleus repulsion, and the electron-electron repulsion. This equation may be further simplified through the Born-Oppenheimer approximation, in which it is assumed that nuclei move on a vastly different timescale than electrons; that is, it is assumed that the nuclei are fixed in space. This removes the $\hat{T}_N$ term of the Hamiltonian and changes the nucleus-nucleus repulsion term to a constant that can be added to the total energy of the system afterwards, leaving the following Hamiltonian to be used in the time-independent Schrödinger equation:

$$\hat{H} = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee} \tag{1.6}$$

Or, more explicitly,

$$\hat{H} = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_i \sum_A \frac{Z_A}{|R_A - r_i|} + \sum_{i<j} \frac{1}{r_{ij}} \tag{1.7}$$

Exact solutions of the time-independent Schrödinger equation are solvable for systems with a single electron, such as the hydrogen atom. The unfortunate truth is that for any chemically relevant system, more than one electron will almost certainly be present, and an exact solution under the given assumptions cannot be determined. Herein lies the interesting problems that both intrigue and plague quantum chemists. Thus, further strategies must be employed to approximate the solutions to these analytically unsolvable problems.

1.2 Hartree-Fock Theory

The wavefunction, as previously seen in Equation 1.1, is a mathematical representation of the electrons in a system. Often, these are built as a product of one-electron molecular orbitals. All functions which make up this product are collectively called the "basis set". In order to satisfy the spin statistical rules of fermions, the wavefunction is required to be antisymmetric with respect to the interchange of electrons. This can be realized by writing it as a linear

combination of atomic orbitals, which can be further condensed into what is known as a Slater determinant:

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1^1 & \psi_2^1 & \psi_3^1 & ... & \psi_N^1 \\ \psi_1^2 & \psi_2^2 & \psi_3^2 & ... & \psi_N^2 \\ \psi_1^3 & \psi_2^3 & \psi_3^3 & ... & \psi_N^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_1^N & \psi_2^N & \psi_3^N & ... & \psi_N^N \end{vmatrix} \tag{1.8}$$

Hartree-Fock (HF) Theory is the most straightforward manner of solving the time-independent Schrödinger equation. In Hartree-Fock Theory, it is assumed that the wavefunction may be written as a single Slater determinant. Through this, it is inherently assumed that electron correlation of a system is averaged over all its electrons.

Although the mean-field treatment of electrons employed by Hartree-Fock theory greatly simplifies the solving process of the Schrödinger equation, the solution itself is often reliable for qualitative results only, even though it often captures almost 99% of the energy of a system. Thus, more rigorous treatments of electron correlation are required to obtain any accurate quantitative results.

1.3 Treatments for Electron Correlation

As stated previously, Hartree-Fock Theory is a mean-field method, and therefore it ignores any electron correlation: the instantaneous repulsion of electrons in a system. Any energy not accounted for by Hartree-Fock is therefore defined as the correlation energy:

$$E_{corr} \equiv E_{exact} - E_{HF} \tag{1.9}$$

Although the correlation energy only constitutes approximately 1% of the total energy of a system, it is considered to be the most important component of a system's total energy

and is the main focus of electronic structure theory due to its importance in replicating and predicting experimental data.

1.3.1 Configuration Interaction Theory

The Hartree-Fock wavefunction, which is by definition single-reference, may be expanded to include all possible excited wavefunctions. This is referred to as the full configuration interaction (FCI) wavefunction:

$$|\Psi_{CI}\rangle = c_0 |\Phi_0\rangle + \left(\frac{1}{1!}\right)^2 \sum_{ia} c_i^a |\Phi_i^a\rangle + \left(\frac{1}{2!}\right)^2 \sum_{\substack{ij \\ ab}} c_{ij}^{ab} \left|\Phi_{ij}^{ab}\right\rangle + \left(\frac{1}{3!}\right)^2 \sum_{\substack{ijk \\ abc}} c_{ijk}^{abc} \left|\Phi_{ijk}^{abc}\right\rangle + ... \quad (1.10)$$

where $i, j, k...$ refer to occupied orbitals and $a, b, c...$ refer to unoccupied orbitals, therefore showing the excitation of electrons from orbital i to orbital a. In this form, each term refers to the number of excitations to the reference wavefunction. So, the wavefunction may be rewritten more compactly as:

$$|\Psi_{CI}\rangle = (c_0 + \hat{C}_1 + \hat{C}_2 + \hat{C}_3 + ... + \hat{C}_N) |\Phi\rangle \quad (1.11)$$

where $\hat{C}_n$ represents the operator for the excitation of an n number of electrons in a CI wavefunction, and N is the total number of electrons in the system.

Using an infinite basis set, energies computed with FCI are exact. However, even using moderately-sized finite basis sets, it is nearly impossible to compute these energies. So, the wavefunction is often truncated after a certain number of excitations. A commonly used truncation is CISD, or configuration interaction with single and double excitations, in which the wavefunction is cut off after $\hat{C}_2$. This significantly decreases the computation time while still retaining most of the correlation energy. However, by truncating the wavefunction, it is no longer size-consistent nor size-extensive; that is, the energy computed with a truncated CI wavefunction on two non-interacting system systems is not equal to the energies of

the two systems computed separately, and the method does not scale linearly with the number of electrons.[3] This problem can be avoided by using a slightly different ansatz of the wavefunction.

1.3.2 Coupled Cluster Theory

Coupled cluster (CC) theory uses a different form of the CI wavefunction that maintains the properties that configuration interaction loses after truncating the wavefunction: the previously discussed size consistency and size extensivity. This is achieved through an exponential ansatz of the wavefunction:

$$|\Psi\rangle = e^{(\hat{T}_1+\hat{T}_2+\hat{T}_3+\ldots)}|\Phi\rangle \tag{1.12}$$

where $\hat{T}_n$ represents the operator for the excitation of an n number of electrons in a CC wavefunction. The non-truncated coupled cluster wavefunction is equivalent to the FCI wavefunction. Therefore, similarly to that of configuration interaction, truncation is necessary to compute energies for chemically relevant systems with any moderately-sized basis set. However, unlike truncated CI methods, truncated coupled cluster maintains the properties of size consistency and size extensivity present in FCI. Because of this, truncated coupled cluster methods are some of the most common *ab initio* methods on small chemical systems due to their high accuracy.

1.3.3 Perturbation Theory

Electron correlation can also be treated with perturbation theory, in which some known reference system is "perturbed" slightly, together forming a more comprehensive solution to the given problem.

In Møller-Plesset Perturbation Theory, the known reference system is the Hartree-Fock solution, and the perturbation is a Taylor series expansion, truncated at the n^{th} order. At higher orders, these methods can diverge while still increasing substantially in cost.[4] Because of this, the only commonly used form of Møller-Plesset Perturbation Theory is its second order form, MP2:

$$E_0^{(2)} = \frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab\rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{1.13}$$

In coupled cluster theory, one can also approximate the solutions including the next order of excitations using many-body perturbation theory arguments.[5,6] However, instead of using Hartree-Fock as the reference as in Møller-Plesset Perturbation Theory, one can use the previous truncation of coupled cluster as the reference. So, to approximate triple excitations with coupled cluster theory [i.e. CCSD(T)], one can use CCSD as the reference, then slightly perturb the system using similar arguments to Møller-Plesset Perturbation Theory to obtain the approximate contributions of triple excitations to the energy.

1.4 Focal Point Approach

The exact energy of a system can be obtained only if the full configuration interaction energy of a system is computed at the complete basis set limit. For all chemically relevant systems, this is intractable; as the size of the basis set increases, more accurate methods (e.g. high-order coupled cluster) become significantly more computationally expensive compared to less accurate methods (e.g. Hartree-Fock). Composite methods have been developed to approximate the energies computed at the complete basis set limit while still including the correlation energy of higher-order quantum chemical methods. One such method is the focal point approach (FPA)[7–10].

The Focal Point Approach is a general means by which the sought FCI/CBS energy may be estimated systematically. Unlike other popular composite methods, such as HEAT[11] and W4[12] which have a set protocol to estimate the same energy, the Focal Point Approach has non-specific guidelines which can be followed to achieve the same goal with the flexibility to increase or decrease the accuracy and coast of utilized methods as necessary. In the Focal Point Approach, energies are computed at various methods using the Dunning correlated consistent (i.e. cc-pVXZ) family of basis sets, which were designed specifically to converge toward the complete basis set limit systematically as the cardinality of each basis set increases. Molecular energies are computed starting with Hartree-Fock and MP2, then with all coupled cluster methods that are computationally feasible for the given system, alternating between those with perturbative excitations and those without (e.g. CCSD, CCSD(T), CCSDT, CCSDT(Q), etc.). At each level of theory, energies are computed with the largest Dunning basis set computationally possible as well as with every basis set smaller than it.

In order to approximate energies at the CBS limit, extrapolation schemes are used. Hartree-Fock energies are extrapolated using a three-point scheme[13,14], as shown below:

$$E_{\mathrm{HF}} = A + Be^{-CX} \tag{1.14}$$

Because correlated methods and Hartree-Fock converge differently, a separate two-point extrapolation scheme[15] is used:

$$E_{\mathrm{corr}} = A + BX^{-3} \tag{1.15}$$

Due to the unreliability of energies computed with double-ζ basis sets, they are generally excluded from use in the energy extrapolations. Therefore, it is necessary, at minimum, to compute Hartree-Fock energies at least up to a pentuple-ζ basis set and correlation

energies at least up to a quadruple-ζ basis set in order to obtain accurate extrapolated energies using the three-point and two-point extrapolation schemes, respectively. However, it is still recommended to compute energies with all methods up to the largest basis set possible so that the most accurate estimate of the energies at the complete basis set limit may be calculated. For correlated methods in which quadruple-ζ quality energies are not computationally feasible, the energy computed with the largest basis set possible can be effectively treated as the extrapolated energy.

Once all extrapolations are complete, the energies can be organized into an incremented focal point table. An example of an incremented focal point table is shown below:

Table 1.1: Incremented focal point table for the enthalpy of formation of formaldehyde oxide (CH_2OO). All energies are given in kcal mol^{-1} relative to the isolated reactants.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	$+\delta$(P)	ΔE_e NET
cc-pVDZ	47.75	–6.53	–6.88	–4.76	–0.19	–1.39	+0.68	–0.05	[+28.63]
cc-pVTZ	46.41	–6.78	–5.27	–4.93	–0.10	[–1.39]	[+0.68]	[–0.05]	[+28.57]
cc-pVQZ	46.17	–6.82	–4.56	–5.00	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.94]
cc-pV5Z	45.84	–7.06	–4.23	–5.05	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.63]
cc-pV6Z	45.77	–7.16	–4.12	–5.08	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.55]
CBS	[45.75]	[–7.29]	[–3.97]	[–5.11]	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.52]

In the table, basis sets are listed in the left-most column with increasing cardinality going down, and methods used to compute the energies are listed in the top row, with increasing order going to the right. Each computed energy, besides those at the Hartree-Fock level of theory, is shown as the difference between it and the energy computed one level of theory less, i.e. the energy to its left in the table. Any energy not explicitly computed is shown in brackets, and the extrapolated energies are shown in the bottom row. Each row of energies, including extrapolated energies, are then summed to show the final energy at each basis set. So, the final FCI/CBS energy is shown in the most bottom-right entry in the table.

1.5 Optimized Virtual Orbitals

Higher-order correlated methods, such as CCSDTQ, scale quite poorly with system size, so the use of these methods on systems with more than a few heavy atoms can quickly become intractable. This can be attributed, in part, to the large number of orbitals which are unoccupied in the ground state, known as virtual orbitals, of which the highest-lying orbitals do not individually contribute substantially to the final energy of the system. Optimized virtual orbitals are an effective method to reduce the size of the virtual orbital space while still retaining the accuracy associated with including the full virtual orbital space.[16,17] This is accomplished by using an invariant unitary rotation matrix onto a subset of the virtual orbital subspace. The contributions of the orbitals on which the unitary rotation was applied are then projected onto the rest of the virtual orbital space, and the orbitals subjected to the rotation are omitted from the subsequent computations. The size of the virtual space retained can be decreased significantly, often by a factor of 2, while still conserving a majority of the correlation energy projected from the omitted orbitals. This can decrease the computational cost of higher-order coupled cluster methods by an order of magnitude or more, allowing for these higher-order methods to be used on larger systems without losing much accuracy.

Chapter 2

Enthalpies of Formation for Criegee Intermediates: A Correlation Energy Convergence Study[1]

[1]Reprinted from J. M. Begley, G. J. R. Aroeira, J. M. Turney, G. E. Douberly and H. F. Schaefer III, *J. Chem. Phys.*, 2023, **158**, 034302, with the permission of AIP Publishing.

2.1 Abstract

Criegee intermediates, formed from the ozonolysis of alkenes, are known to have a role in atmospheric chemistry, including the modulation of the oxidizing capacity of the troposphere. Although studies have been conducted since their discovery, the synthesis of these species in the laboratory has ushered in a new wave of investigations of these structures, both theoretically and experimentally. In some of these theoretical studies, high-order corrections for correlation energy are included to account for the mid multi-reference character found in these systems. Many of these studies include a focus on kinetics; therefore calculated energies should be accurate (<1 kcal/mol in error). In this research, we compute the enthalpies of formation for a small set of Criegee intermediates, including higher-order coupled cluster corrections for correlation energy up to CCSDTQ(P). The enthalpies of formation for formaldehyde oxide, *anti*–acetaldehyde oxide, *syn*–acetaldehyde oxide, and acetone oxide are presented at 0 K as 26.5 kcal mol^{-1}, 15.6 kcal mol^{-1}, 12.2 kcal mol^{-1}, and 0.1 kcal mol^{-1}, respectively. Additionally, we do not recommend the CCSDT(Q) energy correction, as it is approximately twice as large as that of the full CCSDTQ. Half of the CCSDT(Q) energy correction may be included as a reliable, cost-effective estimation of CCSDTQ energies for Criegee intermediates.

2.2 Introduction

Carbonyl oxides, or Criegee intermediates (CIs), were first proposed as intermediate species formed during the ozonolysis of unsaturated hydrocarbons.[18,19] Recently, there has been a surge of interest in Criegee intermediates due to their potential role in atmospheric and combustion chemistry.[20] The unimolecular decomposition of the vibrationally excited CIs, termed "hot" CIs, is a source of OH and HO_2 radicals in the atmosphere.[21–26] The production of these radicals by CIs, in turn, help to modulate the oxidizing capacity of the atmosphere.

Ngyuen and coworkers proposed that a significant fraction of hot CIs (approximately 42% for formaldehyde oxide) can be stabilized through collisional quenching.[27] Such stabilized Criegee intermediates (sCI) are subsequently able to undergo bimolecular reactions, such as those with H_2O,[28–33] NH_3,[34–37] NO_2,[33,38,39] SO_2,[33,40–42] aldehydes,[33,43,44] alcohols,[45–47] and carboxylic acids.[48–50]

Bimolecular reactions involving CIs may affect the atmosphere in different ways. For example, the removal of CIs in the atmosphere is dominated by the reaction with water molecules.[51] The major product of the reaction between the simplest CI, CH_2OO, and water is hydroxymethyl hydroperoxide (HMHP). HMHP has numerous environmental implications which include toxicity to plants and enzymes; it can also decompose into formic acid, a component of acid rain.[20,32,52,53] Another relevant reaction pathway is with SO_2, which can be oxidized by CH_2OO to produce SO_3 that can readily react with water to produce H_2SO_4, although this pathway is considered to be minor in comparison to that of H_2O.[51,54–57]

The simplest CI, formaldehyde oxide (CH_2OO), was first observed experimentally by Taatjes et al. as a product of the oxidation of dimethyl sulfoxide, DMSO.[58] Following that, a new pathway for the formation of CIs was presented where the reaction of CH_2I_2 and O_2 was employed.[54] Various other, larger CIs began to be produced in an analagous manner, such as CH_3CHOO from CH_3CHI_2 and $(CH_3)_2COO$ from $(CH_3)_2Cl_2$.[40,49,55,59–65] These new pathways allowed kinetic, spectroscopic, and barrier height studies to be conducted experimentally, opening the doors for a surge of research into CIs.[28,55,66,67]

For theoretical studies involving kinetics and barrier heights, accurate energy computations are required. Therefore, composite methods including higher-order corrections such as CCSDT(Q) are frequently employed. However, Matthews found that with respect to CCSDTQ computations on CIs, CCSDT(Q) computations had a larger statistical error for equilibrium rotational constants, bond lengths, bond angles, dihedral angles, harmonic frequencies, and zero-point vibrational energies.[68] Moreover, they show that absolute energies

computed with CCSDT(Q) differ from those computed with CCSDTQ by more than 1 kcal mol^{-1}. This phenomenon can be explained by the mid multi-reference character of CIs.

Theoretical enthalpy of formation studies on Criegee related systems have also been conducted. The enthalpy of formation of formaldehyde oxide has been computed in several studies. However, its value is heavily contested, with the enthalpy of formation at 0 K ranging from 23.8 kcal mol^{-1} to 31.7 kcal mol^{-1}.[69–72] Additionally, Kettner et al. computed the enthalpies of formation of both the *anti*- and *syn*- conformers of acetaldehyde oxide with W3-F12 theory, which uses coupled cluster methods up to CCSDT(Q).[73,74]

In this research, we compute accurate enthalpies of formation for the four CIs presented (see Figure 2.1) using high order coupled cluster to assess convergence. The CIs studied are formaldehyde oxide [CH_2OO], both the *anti* and *syn* conformers of acetaldehyde oxide [CH_3CHOO], and acetone oxide [$(CH_3)_2COO$].

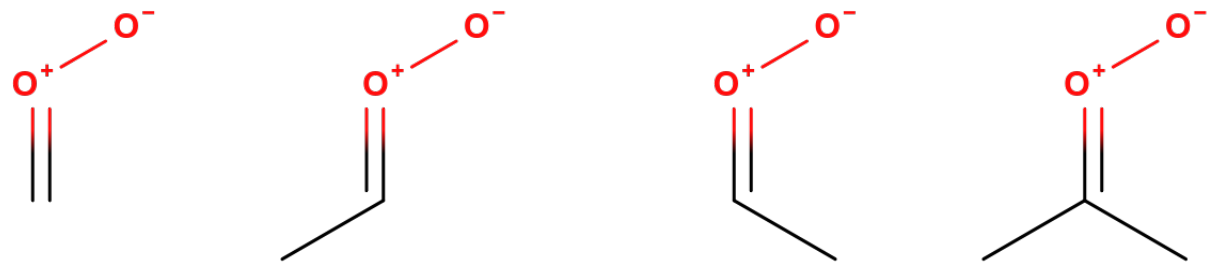

Figure 2.1: Structures of the four simplest Criegee intermediates. From left to right, formaldehyde oxide, *syn*-acetaldehyde oxide, *anti*-acetaldehyde oxide, and acetone oxide.

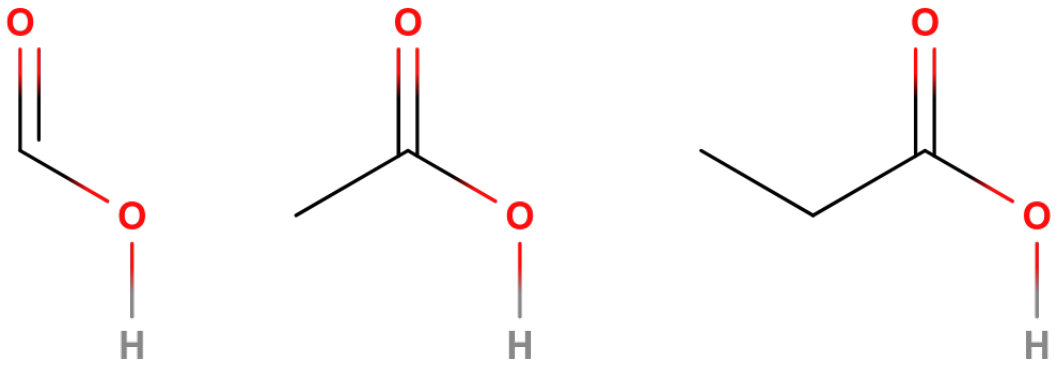

Figure 2.2: Structures of the three simplest carboxylic acids. From left to right, formic acid, acetic acid, and propanoic acid.

2.3 Methods

The enthalpy for each reaction shown in Eq. (2.1)-Eq. (2.4) was computed in a gaseous, 0 K environment.

$$\mathrm{HCOOH_{(g)}} \rightarrow \mathrm{CH_2OO_{(g)}} \tag{2.1}$$

$$\mathrm{CH_3COOH_{(g)}} \rightarrow \mathit{anti}-\mathrm{CH_3CHOO_{(g)}} \tag{2.2}$$

$$\mathrm{CH_3COOH_{(g)}} \rightarrow \mathit{syn}-\mathrm{CH_3CHOO_{(g)}} \tag{2.3}$$

$$\mathrm{CH_3CH_2COOH_{(g)}} \rightarrow \mathrm{(CH_3)_2COO_{(g)}} \tag{2.4}$$

The enthalpy of formation, $\Delta_R H$, of each CI was then determined from each reaction using the enthalpy of formation of the corresponding constitutionally isomeric carboxylic acid (see Fig. 2.2), as shown in Eq. (2.5). The enthalpy of formation for each carboxylic acid

was taken from the Active Thermochemical Tables (ATcT) database of Ruscic.[75]

$$\Delta_f H = \Delta_R H + \Delta_f H_{\mathrm{ROOH}} \tag{2.5}$$

Geometries and harmonic vibrational frequencies of all species were obtained using coupled cluster theory with single, double, and perturbative triple excitations [CCSD(T)], available from the CFOUR 2.0 quantum chemistry package.[76] These geometry optimizations and harmonic vibrational frequency computations utilized the atomic natural orbital (ANO) family of basis sets.[77,78] All geometry optimizations and vibrational frequency computations utilized the large ANO2 basis set (H:[4s3p2d1f]; C,O:[5s4p3d2f1g]) except for acetone oxide and propanoic acid, which used the ANO1 basis set (H:[4s2p1d]; C,O:[4s3p2d1f]).

To obtain reliable energies, the focal point analysis (FPA) technique developed by Allen and coworkers[7,9,79] was used. Hartree–Fock energies[80,81] were computed using the Dunning correlation consistent basis sets,[82] cc-pVXZ (X = D, T, Q, 5, 6). Correlation energies were obtained using second-order Møller–Plesset perturbation theory (MP2)[83], coupled cluster[5] with single and double excitations (CCSD)[84–87], and with perturbative triple excitations [CCSD(T)][88–90] with the cc-pVXZ basis sets (X = D, T, Q, 5). Additionally, correlation energies using coupled cluster with full triple excitations [CCSDT][91–93] were computed with the cc-pVXZ basis sets (X = D, T), while energies using coupled cluster with perturbative quadruple excitations [CCSDT(Q)][94,95] and with full quadruple excitations [CCSDTQ][96–99] were computed using the cc-pVDZ basis set as well as with the cc-pVTZ basis set for formaldehyde oxide and formic acid. For formaldehyde oxide and formic acid only, correlation energies were also computed using coupled cluster with perturbative quintuple excitations [CCSDTQ(P)][100,101] with the cc-pVDZ basis set as well as up to CCSD(T) with the cc-pV6Z basis set. Energies are presented in kcal mol^{-1}, with the thermochemical definition of the calorie used. (1 cal = 4.184 J)

Geometry optimizations and energy computations were carried out using the frozen core approximation. All energy computations were performed using the MRCC quantum chemistry program.[102,103]

Energies were extrapolated to the complete basis set (CBS) limit using a three-point formula shown in Eq. (2.6) for Hartree-Fock energies.[13,14]

$$E = A + Be^{-CX}. \tag{2.6}$$

Additionally, the two-point formula shown in Eq. (2.7) was used for the extrapolation of correlation energies up to CCSD(T).[15]

$$E_{\text{corr}} = A + BX^{-3} \tag{2.7}$$

Higher-order additive corrections for correlation energy include CCSDT/cc-pVTZ, CCSDT(Q)/cc-pVDZ, and CCSDTQ/cc-pVDZ. Additionally, CCSDTQ(P)/cc-pVDZ energy computations are included for the formaldehyde oxide system. Because the CCSDTQ(P)/cc-pVDZ energy computation is highly expensive, it was feasible for the formaldehyde oxide system only. A final additive correction, Δ, is computed as shown in Equation 2.8. All corrections were computed using the CFOUR 2.0 quantum chemistry package and are further explained below.

$$\Delta = \Delta_{\text{DBOC}} + \Delta_{\text{core}} + \Delta_{\text{relativistic}} + \Delta_{\text{ZPVE}} \tag{2.8}$$

The diagonal Born–Oppenheimer correction (DBOC) was computed at the CCSD level using the ANO0 basis set. The frozen-core correction was computed at the CCSD(T) level using the weighted core-valence cc-pwCVTZ basis set[104] in two computations. An energy computation was obtained with a frozen core approximation (FC) and with all electrons (AE) correlated.

Table 2.1: Incremented focal point table for the enthalpy of formation of formaldehyde oxide (CH_2OO). All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ(P)/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\text{FPA}} + \Delta_{\text{DBOC}} + \Delta_{\text{core}} + \Delta_{\text{relativistic}} + \Delta_{\text{ZPVE}}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	$+\delta$(P)	ΔE_e NET
cc-pVDZ	47.75	–6.53	–6.88	–4.76	–0.19	–1.39	+0.68	–0.05	[+28.63]
cc-pVTZ	46.41	–6.78	–5.27	–4.93	–0.10	[–1.39]	[+0.68]	[–0.05]	[+28.57]
cc-pVQZ	46.17	–6.82	–4.56	–5.00	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.94]
cc-pV5Z	45.84	–7.06	–4.23	–5.05	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.63]
cc-pV6Z	45.77	–7.16	–4.12	–5.08	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.55]
CBS	[45.75]	[–7.29]	[–3.97]	[–5.11]	[–0.10]	[–1.39]	[+0.68]	[–0.05]	[+28.52]

CCSDTQ(P)/CBS+Δ: Δ_fH = 28.52 + 0.05 + 0.27 + (–0.15) + (–2.01) = 26.67 kcal mol^{-1}

The correction was taken as the difference between the two, as shown in Equation 2.9.

$$\Delta_{\text{core}} = E_{\text{AE-CCSD(T)}} - E_{\text{FC-CCSD(T)}} \tag{2.9}$$

The relativistic effects of the system were determined with an X2C-recontracted cc-pCVTZ basis set.[105–110] Two energy computations were obtained with and without spin-free exact two-component theory at the one electron level (X2C–1e). The relativistic correction was taken as the difference between the two as shown in Equation 2.10.

$$\Delta_{\text{relativistic}} = E_{\text{AE-CCSD(T)/X2C-1e}} - E_{\text{AE-CCSD(T)}} \tag{2.10}$$

Finally, the zero-point vibrational energy (ZPVE) was obtained from the CCSD(T)/ANO2 harmonic vibrational frequency computation.

In order to reduce the cost of higher order coupled cluster computations, the optimized virtual orbitals (OVOs) method was utilized to truncate the size of the virtual space for systems larger than formaldehyde oxide.[16] For both conformers of acetaldehyde oxide and acetic acid, 75% of the orbital virtual space was retained, while for acetone oxide and propanoic acid, 50% of the orbital virtual space was retained for each respective molecule.

Table 2.2: Errors for the enthalpy of formation (kcal mol^{-1}) of formaldehyde oxide computed with various methods compared to CCSDTQ(P)/CBS+Δ results and CCSDTQ/CBS+Δ results.

Method	Error compared to CCSDTQ	Error compared to CCSDTQ(P)
CCSD(T)/CBS+Δ	0.34	0.39
75% - CCSD(T)/CBS+Δ	0.84	0.89
50% - CCSD(T)/CBS+Δ	0.73	0.78
CCSDT(Q)/CBS+Δ	–0.29	–0.24
75% - CCSDT(Q)/CBS+Δ	–0.54	–0.59
50% - CCSDT(Q)/CBS+Δ	–0.65	–0.70
CCSDTQ/CBS+Δ	0.00	0.05
75% - CCSDTQ/CBS+Δ	0.03	0.08
50% - CCSDTQ/CBS+Δ	–0.08	–0.03
CCSDTQ(P)/CBS+Δ	–0.05	0.00
75% - CCSDTQ(P)/CBS+Δ	–0.01	0.04
50% - CCSDTQ(P)/CBS+Δ	–0.12	–0.07
Nguyen et al.[69]	1.38	1.43
Harding et al.[70]	2.38	2.43
Olzmann et al.[71]	–2.92	–2.87
Cremer et al.[72]	4.98	5.03
ATcT	–0.12	–0.07

These percentages were the largest computationally feasible proportions (with our resources) of virtual space retained. To assess the error introduced due to this approximation, the enthalpy of formation of formaldehyde oxide was also computed using the OVOs approach.

2.4 Results and Discussion

2.4.1 Formaldehyde Oxide

Canonical Orbitals

The incremented focal point table for the formation of formaldehyde oxide (CH_2OO) along with the enthalpy of formation is shown in Table 2.1. The enthalpy of formation for this system is computed to be 26.67 kcal mol^{-1} at the CCSDTQ(P)/CBS+Δ level of theory.

Table 2.3: Incremented focal point table for the enthalpy of formation of formaldehyde oxide (CH_2OO) *with 75% of the optimized virtual orbital space retained*. All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ(P)/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\text{FPA}} + \Delta_{\text{DBOC}} + \Delta_{\text{core}} + \Delta_{\text{relativistic}} + \Delta_{\text{ZPVE}}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	$+\delta$(P)	ΔE_e NET
cc-pVDZ	47.75	–7.11	–6.91	–4.77	–0.22	–1.19	+0.57	–0.07	[+28.04]
cc-pVTZ	46.41	–7.17	–5.28	–4.89	–0.11	[–1.19]	[+0.57]	[–0.07]	[+28.28]
cc-pVQZ	46.17	–6.83	–4.54	–4.96	[–0.11]	[–1.19]	[+0.57]	[–0.07]	[+29.04]
cc-pV5Z	45.84	–7.07	–4.23	–5.03	[–0.11]	[–1.19]	[+0.57]	[–0.07]	[+28.70]
cc-pV6Z	45.77	[–7.07]	[–4.23]	[–5.03]	[–0.11]	[–1.19]	[+0.57]	[–0.07]	[+28.63]
CBS	[45.59]	[–7.32]	[–3.91]	[–5.12]	[–0.11]	[–1.19]	[+0.57]	[–0.07]	[+28.56]

CCSDTQ(P)/CBS+Δ: Δ_fH = 28.56 + 0.05 + 0.27 + (–0.15) + (–2.01) = 26.71 kcal mol^{-1}

Table 2.4: Incremented focal point table for the enthalpy of formation of formaldehyde oxide (CH_2OO) *with 50% of the optimized virtual orbital space retained* including the contributions of CCSDT(Q)/cc-pVTZ and and CCSDTQ/cc-pVTZ. All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ(P)/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\text{FPA}} + \Delta_{\text{DBOC}} + \Delta_{\text{core}} + \Delta_{\text{relativistic}} + \Delta_{\text{ZPVE}}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	$+\delta$(P)	ΔE_e NET
cc-pVDZ	47.75	+9.15	–5.76	–4.07	–0.26	–0.54	+0.26	–0.05	[+46.48]
cc-pVTZ	46.41	–5.17	–5.22	–4.31	–0.15	–1.15	+0.48	[–0.05]	[+30.84]
cc-pVQZ	46.17	–6.45	–4.57	–4.79	[–0.15]	[–1.15]	[+0.48]	[–0.05]	[+29.49]
cc-pV5Z	45.84	–6.89	–4.29	–4.95	[–0.15]	[–1.15]	[+0.48]	[–0.05]	[+28.84]
cc-pV6Z	45.77	[–6.89]	[–4.29]	[–4.95]	[–0.15]	[–1.15]	[+0.48]	[–0.05]	[+28.77]
CBS	[45.75]	[–7.35]	[–3.98]	[–5.12]	[–0.15]	[–1.15]	[+0.48]	[–0.05]	[+28.45]

CCSDTQ(P)/CBS+Δ: Δ_fH = 28.45 + 0.05 + 0.27 + (–0.15) + (–2.01) = 26.60 kcal mol^{-1}

The absolute error of the energy computation with respect to the data from ATcT is −0.07 kcal mol^{-1}. From the focal point table, it can be seen that the CCSDT(Q) contribution is twice as large in magnitude than that from CCSDTQ. The relatively small contribution of perturbative quintuple excitations suggests the correlation energy is becoming well converged at the CCSDTQ level of theory.

The computed enthalpy of formation at the CCSD(T)/CBS+Δ level of theory is 27.53 kcal mol^{-1}. With respect to the experimental data from ATcT, the absolute error is +0.79 kcal mol^{-1}. The magnitude of the absolute error of the determined enthalpy of formation of fomaldehyde oxide at the CCSD(T)/CBS+Δ level of theory is an order of magnitude larger than that at the CCSDTQ(P)/CBS+Δ level of theory. This suggests that higher–excitation corrections are necessary for highly accurate energy computations for CIs.

The absolute errors of enthalpy of formation computations without higher-order energy corrections are shown in Table 2.2. With respect to the computed CCSDTQ(P)/CBS+Δ enthalpy of formation, the absolute error of the CCSD(T)/CBS+Δ enthalpy of formation is +0.86 kcal mol^{-1}. The magnitude of the absolute error to the enthalpy of formation predicted at the CCSD(T)/CBS+Δ level of theory is an order of magnitude larger than that at the CCSDTQ(P)/CBS+Δ level of theory. As previously suggested, higher-order corrections are therefore necessary for accurate enthalpy of formation computations for CIs. However, the computed enthalpy of formation at the CCSDT(Q)/CBS+Δ level of theory is 26.04 kcal mol^{-1}. The absolute error of this computation compared to the CCSDTQ(P) enthalpy of formation is −0.63 kcal mol^{-1}. Although this is more accurate than the computed CCSD(T)/CBS+Δ enthalpy of formation, it is still not entirely converged at this level of theory.

Absolute errors of various literature values compared to the computed CCSDTQ(P)/CBS+Δ enthalpy of formation are also shown in Table 2.2. All compared literature values have absolute errors of greater than 1 kcal mol^{-1}.[69–72] Given that our com-

Table 2.5: Incremented focal point table for the enthalpy formation of anti-acetaldehyde oxide (anti$-CH_3COOH$). 75% of the optimized virtual orbital space for both molecules has been retained. All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\mathrm{FPA}} + \Delta_{\mathrm{DBOC}} + \Delta_{\mathrm{core}} + \Delta_{\mathrm{relativistic}} + \Delta_{\mathrm{ZPVE}}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	ΔE_e NET
cc-pVDZ	31.02	–2.65	–6.99	–3.66	–0.22	–1.06	+0.54	[+16.97]
cc-pVTZ	30.15	–2.89	–5.60	–3.85	–0.16	[–1.06]	[+0.54]	[+17.13]
cc-pVQZ	30.13	–2.80	–5.01	–3.97	[–0.16]	[–1.06]	[+0.54]	[+17.66]
cc-pV5Z	29.80	–3.03	–4.73	–4.04	[–0.16]	[–1.06]	[+0.54]	[+17.32]
cc-pV6Z	29.74	[–3.03]	[–4.73]	[–4.04]	[–0.16]	[–1.06]	[+0.54]	[+17.25]
CBS	[29.72]	[–3.27]	[–4.44]	[–4.11]	[–0.16]	[–1.06]	[+0.54]	[+17.23]

CCSDTQ/CBS+Δ: Δ_fH = 17.23 + 0.06 + 0.26 + (–0.16) + (–1.77) = 15.62 kcal mol^{-1}

puted enthalpy of formation is similar to that shown in the ATcT, we conclude that the familiar methods of computing this value in the literature are not adequate for accurate enthalpy of formation calculations.

Although the harmonic zero-point vibrational energy is cheaper to compute than its anharmonic counterpart, the accuracy of this assumption should be tested for highly accurate energy computations. The zero-point vibrational energy correction for the formation of formaldehyde oxide with the anharmonic correction was computed to be -2.04 kcal mol^{-1}. This is 0.03 kcal mol^{-1} lower than the determined harmonic zero-point vibrational energy of -2.01 kcal mol^{-1}. Because of this small difference, we conclude that the use of the harmonic zero-point vibrational energy may be sufficient for the determination of energies of larger Criegee intermediates.

Optimized Virtual Orbitals

The incremented focal point table for the formation of formaldehyde oxide with 75% of the virtual orbital space retained is shown in Table 2.3, along with the enthalpy of formation. The computed enthalpy of formation with the reduced virtual orbital space is 26.71 kcal mol^{-1}.

Table 2.6: Incremented focal point table for the enthalpy of formation of syn-acetaldehyde oxide (syn−CH_3COOH). 75% of the optimized virtual orbital space for both molecules has been retained. All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\mathrm{FPA}} + \Delta_{\mathrm{DBOC}} + \Delta_{\mathrm{core}} + \Delta_{\mathrm{relativistic}} + \Delta_{\mathrm{ZPVE}}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	ΔE_e NET
cc-pVDZ	29.58	–5.26	–6.25	–4.03	–0.16	–1.07	+0.53	[+13.34]
cc-pVTZ	29.15	–6.08	–4.71	–4.35	–0.07	[–1.07]	[+0.53]	[+13.39]
cc-pVQZ	29.26	–6.18	–4.07	–4.50	[–0.07]	[–1.07]	[+0.53]	[+13.89]
cc-pV5Z	29.03	–6.40	–3.81	–4.58	[–0.07]	[–1.07]	[+0.53]	[+13.63]
cc-pV6Z	28.99	[–6.40]	[–3.81]	[–4.58]	[–0.07]	[–1.07]	[+0.53]	[+13.59]
CBS	[28.98]	[–6.63]	[–3.54]	[–4.66]	[–0.07]	[–1.07]	[+0.53]	[+13.54]

CCSDTQ/CBS+Δ: Δ_fH = 13.54 + 0.05 + 0.25 + (–0.16) + (–1.53) = 12.15 kcal mol^{-1}

With respect to the canonical computation, the absolute error is +0.03 kcal mol^{-1}. Because the absolute error is within 0.25 kcal mol^{-1}, we conclude that the use of optimized virtual orbitals up to a reduction of 75% is a reasonable computation time reduction strategy for larger systems of this type. Due to the reduction of the virtual space, the relative CCSDT(Q) correction has been reduced considerably. However, it is still approximately twice as large as the contribution of CCSDTQ.

The incremented focal point table for the formation of formaldehyde oxide with 50% of the virtual orbital space retained is shown in Table 2.4, along with the enthalpy of formation. To explore the basis set dependence of the CCSDT(Q) contribution, additional energy computations for formaldehyde oxide and formic acid were computed at the CCSDT(Q)/cc-pVTZ and CCSDTQ/cc-pVTZ levels of theory. The computed enthalpy of formation with the reduced virtual orbital space is 26.60 kcal mol^{-1}. With respect to the canonical computation, the absolute error is −0.08 kcal mol^{-1}. Because the absolute error is within 0.25 kcal mol^{-1}, we conclude that with the use of optimized virtual orbitals, up to a reduction of 50% may be a reasonable computation time reduction strategy for larger systems of this type. Following a similar trend to that of the computation with 75% of the optimized virtual orbital space

retained, the CCSDT(Q) contribution has decreased as the size of the virtual space decreases, but is still approximately twice as large as the contribution of CCSDTQ. Additionally, the CCSDT(Q) correction at the cc-pVTZ level is more than twice as large as that at the cc-pVDZ level and is still approximately twice as large in magnitude as the CCSDTQ correction at this level of theory, indicating that the higher-order corrections have a significant basis set dependence for this class of molecules.

The enthalpies of formation calculated without higher-order energy corrections (i.e. CCSD(T)/CBS+Δ) with both 75% and 50% of the virtual orbitals retained are shown below Tables 2.3 and 2.4, respectively. The errors of these computations as well as the errors of literature values relative to the computed enthalpy of formation with the full set of canonical orbitals are shown in Table 2.2. The computed enthalpy of formation with 75% of the virtual orbital space retained is 27.56 kcal mol^{-1}. With respect to the computed enthalpy of formation with canonical orbitals, the error is +0.89 kcal mol^{-1}. The computed enthalpy of formation with 50% of the virtual orbital space retained is 27.45 kcal mol^{-1}. With respect to the computed enthalpy of formation with canonical orbitals, the error is +0.78 kcal mol^{-1}. Given the magnitude of error, we conclude that higher-order corrections are still necessary for accurate energy computations of CIs with an optimized virtual space.

2.4.2 Acetaldehyde Oxide

The incremented focal point tables for the enthalpy of formation of *anti*–acetaldehyde oxide (*anti*−CH_2COOH) and *syn*–acetaldehyde oxide (*syn*−CH_3COOH) are shown in Tables 2.5 and 2.6, respectively. The computed enthalpy of formation for *anti*–acetaldehyde oxide is 15.62 kcal mol^{-1}, while the computed enthalpy of formation for *syn*–acetaldehyde oxide is 12.15 kcal mol^{-1}. Again, the CCSDT(Q) contribution is approximately twice the magnitude of the CCSDTQ contribution for both conformers. Unlike the enthalpy of formation for formaldehyde oxide, there is no reference enthalpy of formation given in the ATcT[75] for the

Table 2.7: Errors for the enthalpy of formation (kcal mol^{-1}) of *anti*–acetaldehyde oxide computed with various methods compared to canonical CCSDTQ/CBS+Δ data.

Method	Error
CCSD(T)/CBS+Δ	0.68
CCSDT(Q)/CBS+Δ	–0.55
Kettner et al.[73]	0.01

two conformers of acetaldehyde oxide. However, Kettner *et al.* computed the enthalpies of formation of *anti*–acetaldehyde oxide and *syn*–acetaldehyde oxide as 15.63 and 12.26 kcal mol^{-1}, respectively.[73] The absolute errors of these values compared to the CCSDTQ/CBS+Δ enthalpies of formation are shown in Tables 2.7 and 2.8, respectively. For both molecules, the literature values agree with the computed CCSDTQ/CBS+Δ values.

The absolute errors in the enthalpy of formation computations of *anti*–acetaldehyde oxide and of *syn*–acetaldehyde oxide without higher-order energy corrections is shown in Tables 2.7 and 2.8, respectively. The computed CCSD(T)/CBS+Δ enthalpy of formation of *anti*–acetaldehyde oxide is 16.30 kcal mol^{-1}, while the computed CCSD(T)/CBS+Δ enthalpy of formation of *syn*–acetaldehyde oxide is 12.77 kcal mol^{-1}. These computed enthalpies of formation are 0.62 and 0.68 kcal mol^{-1} less than their CCSDTQ/CBS+Δ counterparts. Therefore, we again conclude that higher-order corrections are necessary for truly reliable relative energy computations for CIs. Additionally, the CCSDT(Q)/CBS+Δ enthalpies of formation have absolute errors of -0.54 and -0.53 kcal mol^{-1} compared to CCSDTQ/CBS+Δ. Similar to the formaldehyde oxide computations, CCSDT(Q)/CBS+Δ is more accurate than CCSD(T)/CBS+Δ, but both methods are still not adequate for precise enthalpy of formation computations.

Table 2.8: Errors for the enthalpy of formation (kcal mol^{-1}) of *syn*–acetaldehyde oxide computed with various methods compared to canonical CCSDTQ/CBS+Δ data.

Method	Error
CCSD(T)/CBS+Δ	0.62
CCSDT(Q)/CBS+Δ	–0.53
Kettner et al.[73]	0.11

Table 2.9: Incremented focal point table for the enthalpy of formation of acetone oxide [$(CH_3)_2COO$]. 50% of the optimized virtual orbital space for both molecules has been retained. All energies are given in kcal mol^{-1} relative to the isolated reactants. The final enthalpy of formation at both the CCSDTQ/CBS and the CCSD(T)/CBS levels is presented as $\Delta_f H = E_{\rm FPA} + \Delta_{\rm DBOC} + \Delta_{\rm core} + \Delta_{\rm relativistic} + \Delta_{\rm ZPVE}$ below the table.

Basis Set	ΔE_e HF	$+\delta$MP2	$+\delta$CCSD	$+\delta$(T)	$+\delta$T	$+\delta$(Q)	$+\delta$Q	ΔE_e NET
cc-pVDZ	16.03	–1.75	–4.93	–3.29	–0.13	–0.51	+0.24	[+5.66]
cc-pVTZ	15.10	–2.85	–4.88	–3.42	–0.12	[–0.51]	[+0.24]	[+3.56]
cc-pVQZ	15.11	–3.09	–4.30	–3.70	[–0.12]	[–0.51]	[+0.24]	[+3.63]
cc-pV5Z	14.82	–3.58	–4.06	–3.86	[–0.12]	[–0.51]	[+0.24]	[+2.93]
cc-pV6Z	14.78	[–3.58]	[–4.06]	[–3.86]	[–0.12]	[–0.51]	[+0.24]	[+2.89]
CBS	[14.77]	[–4.09]	[–4.23]	[–3.96]	[–0.12]	[–0.51]	[+0.24]	[+2.10]

CCSDTQ/CBS+Δ: Δ_fH = 2.10 + 0.03 + 0.16 + (–0.17) + (–2.19) = 0.10 kcal mol^{-1}

Table 2.10: Errors for the enthalpy of formation (kcal mol^{-1}) of acetone oxide computed with various methods compared to canonical CCSDTQ/CBS+Δ data.

Method	Error
CCSD(T)/CBS+Δ	0.39
CCSDT(Q)/CBS+Δ	–0.24

2.4.3 Acetone Oxide

The incremented focal point table for the formation of acetone oxide [$(CH_3)_2COO$], along with the enthalpy of formation, is shown in Table 2.9. The enthalpy of formation for this system is computed to be 0.10 kcal mol^{-1}. As with the mono-substituted CIs, the contribution of CCSDT(Q) is twice as large as that of CCSDTQ. Additionally, the CCSDT(Q) contribution with 50% of the retained orbital virtual space is less than 1 kcal mol^{-1}. We again attribute this to the decreased size of the optimized virtual space. Even still, the CCSDT(Q) correction is twice as large as the CCSDTQ correction.

The absolute errors of enthalpy of formation computations of acetone oxide without higher-order energy corrections is shown in Table 2.10. The CCSD(T)/CBS+Δ enthalpy of formation is 0.49 kcal mol^{-1}. With respect to the computed CCSDTQ/CBS+Δ enthalpy of formation, the absolute error of the CCSD(T)/CBS+Δ enthalpy of formation is +0.39 kcal mol^{-1}. Similarly to formaldehyde oxide and acetaldehyde computations, higher-excitation energy corrections are necessary for truly reliable energy computations. The CCSDT(Q)/CBS+Δ enthalpy of formation is -0.14 kcal mol^{-1}. With respect to the computed CCSDTQ/CBS+Δ enthalpy of formation, the absolute error of the CCSDT(Q)/CBS+Δ enthalpy of formation is -0.24 kcal mol^{-1}. These enthalpy of formation computations are still more accurate than the CCSD(T)/CBS+Δ computations. However, the overcorrection of CCSDT(Q)/cc-pVDZ leads to inaccuracies in computations.

2.4.4 Error Analysis

For all computed enthalpies of formation, both the diagonal Born–Oppenheimer correction and the relativistic correction are relatively small in magnitude, and therefore are not considered to be large sources of error. The frozen core approximation is larger than expected (<0.20 kcal mol^{-1}) for all computed enthalpies of formation, but is still within the bounds of reasonable

approximation and is also not considered to be a large source of error. For the enthalpies of formation of formaldehyde oxide and both conformers of acetaldehyde oxide, the basis set extrapolation from cc-pV6Z to the CBS energy is less than -0.10 kcal mol^{-1}, so this is also within the reasonable bounds of error. However, the extrapolation for the enthalpy of formation of acetone oxide is -0.79 kcal mol^{-1}. This larger extrapolation can be attributed to the reduced size of the optimized virtual space, as shown by the increase in the extrapolation energy of formaldehyde oxide as the optimized virtual space is reduced from the canonical (Table 2.1, -0.03 kcal mol^{-1}) to 50% (Table 2.4, -0.32 kcal mol^{-1}).

For all computed enthalpies of formation, the relative zero-point vibrational energy is greater than 1.50 kcal mol^{-1}. In the case of acetone oxide, the contribution from the zero-point vibrational energy (-2.19 kcal mol^{-1}) is larger in magnitude than the enthalpy of reaction (2.10 kcal mol^{-1}). Alternate reactions were explored to alleviate this, but no other reactions explored had smaller zero-point vibrational energy corrections.

The largest uncertainty for these computations likely comes from the higher-excitation additive corrections for correlation energy. For formaldehyde oxide, the only enthalpy of formation computed with a canonical orbital virtual space, the correlation energy contribution of these higher-order additive corrections is -0.86 kcal mol^{-1}. This contribution decreases in magnitude as the size of the molecule increases, but this again is attributed to the decreased size of the optimized virtual space required for these computations. Therefore, we expect this contribution to be of a similar size or larger with all canonical orbitals. Additionally, there is a significant basis set dependence in CCSDT(Q) and CCSDTQ energy computations, as shown in Table 2.4. This basis set dependence could not be further explored with larger systems or with a larger optimized virtual space due to the computational cost of CCSDTQ/cc-pVTZ.

Finally, we estimate the error for the computed enthalpy of formation of all considered molecules through the focal point analysis technique developed by Allen and coworkers[7,9,79]. From this, we present the enthalpies of formations for the considered CIs as 26.49 ± 0.12

kcal mol^{-1} for formaldehyde oxide, 15.62 ± 0.10 kcal mol^{-1} for *anti*–acetaldehyde oxide, 12.15 ± 0.09 kcal mol^{-1} for *syn*–acetaldehyde oxide, and 0.10 ± 0.13 kcal mol^{-1} for acetone oxide. It is important to note that the error bar for the enthalpy of formation of acetone oxide is larger than the enthalpy of formation itself. However, since the magnitude of this error bar is comparable to the others, we conclude that the value for the enthalpy of formation is still accurate.

2.5 Conclusions

In this study, high-level coupled cluster techniques were used to compute the enthalpies of formation of various Criegee intermediates. The geometries for formic acid, acetic acid, formaldehyde oxide, *anti*–acetaldehyde oxide, and *syn*–acetaldehyde oxide were optimized at the CCSD(T)/ANO2 level, while the geometries for acetone oxide and propanoic acid were optimized at the CCSD(T)/ANO1 level. Computed energies for each species were extrapolated to the CBS limit and included corrections for the zero-point vibrational energy, diagonal Born–Oppenheimer correction, the frozen core approximation, relativistic corrections, and higher-order energy corrections. As a cost reduction strategy, energy computations for acetaldehyde oxide conformers and acetic acid were computed with 75% of the optimized virtual orbital space retained, while energy computations for acetone oxide and propanoic acid were computed with 50% of the optimized virtual orbital space retained. Computed enthalpies of formation for formaldehyde oxide, both acetaldehyde oxide conformers, and acetone oxide are reported at these levels.

The results for formaldehyde oxide indicate that the correlation energy for this system appear converged at the CCSDTQ level, as confirmed by the small (0.05 kcal mol^{-1}) CCSDTQ(P) contribution to the CCSDTQ(P)/CBS+Δ result. This CCSDTQ(P)/CBS+Δ result agrees with the enthalpy of formation given in ATcT, with an absolute deviation of 0.07 kcal mol^{-1}. However, there is a large deviation between our results and other theoretical

predictions in the literature, as shown in Table 2.2. Each literature value presented has a deviation of at least 1 kcal mol^{-1} compared to the CCSDTQ(P)/CBS+Δ results. Because the computed enthalpy of formation of formaldehyde oxide with 50% of the optimized virtual orbital space retained has an error of less than 0.10 kcal mol^{-1} compared to the canonical computation, we conclude that this approximation is a feasible cost-saving technique for more computationally expensive computations, such as those of acetaldehyde oxide and acetone oxide.

The quadruples contribution of CCSDT(Q)/cc-pVDZ for all four investigated CIs is approximately twice as large as that of CCSDTQ/cc-pVDZ. A similar trend is shown with the contributions of CCSDT(Q)/cc-pVTZ and CCSDTQ/cc-pVTZ for the formation of formaldehyde oxide with 50% of the optimized virtual space retained. Given this, we conclude that CCSDT(Q) is a questionable high-level additive correction for energy computations involving Criegee intermediates. As a result, we do not recommend the inclusion of CCSDT(Q) in composite approaches for such molecules. Because the CCSDT(Q) additive correction is approximately twice as large as that of CCSDTQ, we suggest the inclusion of one half of the CCSDT(Q) additive correction for this class of molecule as an approximation to the full CCSDTQ results, as methods beyond CCSDT(Q) are generally considered to be too computationally expensive. However, due to the similar chemical motif of the molecules studied, we recommend that one uses judgement before the implementation of this treatment based on the size of the CCSDT(Q) additive correction compared to the CCSDT additive correction.

Chapter 3

Diatomic Molecules at the Highest Levels of Electronic Structure[1]

[1] J. M. Begley, J. M. Turney, and H. F. Schaefer III. To be submitted to a peer-reviewed journal.

3.1 Introduction

The widespread use of coupled cluster theory[5,111,112] has paved the way for countless highly accurate studies in *ab initio* quantum chemistry. In particular, coupled cluster with single, double, and perturbative triple excitations [CCSD(T)][88–90] has received a great deal of praise due to the lower cost compared to configuration interaction methods and relatively high accuracy compared to DFT methods. As such, it has been termed the "gold standard" of computational chemistry. In a similar manner, the more accurate CCSDT(Q)[94,95] method, with its perturbative quadruple excitations, has been termed the "platinum standard".[113,114]

Previous studies have shown that for "well-behaved systems", such as the isometrization between dimethylcarbene and propane, CCSDT(Q) approximates CCSDTQ energies quite well.[68] For the water dimer, it was also shown that the geometry obtained at the CCSD(T) level of theory is very similar to that of CCSDT, and the geometry obtained at the CCSDT(Q) level of theory is very similar to that of CCSDTQ.[115] In terms of diatomic molecules, it has also been shown by Sordo that CCSD(T) can obtain similar dissociation energies and electron affinities to those computed with CCSDT.[116] For these reasons, perturbative excitations have been used frequently for mostly single-reference systems.

The use of perturbative excitations in coupled cluster theory may not prove as accurate when used on multi-reference systems, however. For example, in the Sordo study on diatomic molecules, it was shown that the multi-reference C_2 molecule produced inaccurate dissociation energies and electron affinities with CCSD(T). It was also shown that for a set of 13 multi-reference molecules including XO and XOO radicals, CCSDT was necessary to compute accurate enthalpies of formation and molecular geometries.[117] For Criegee intermediates, it has been shown in both a previous study by the authors as well as one by Matthews that CCSDT(Q) performs quite poorly compared to CCSDTQ when computing enthalpies of formation, total energies, equilibrium rotational constants, geometric parameters, harmonic

frequencies, and harmonic vibrational zero point energies.[68,118] In a study by Hobza and coworkers on dimer systems, CCSDT(Q) energies were consistently less accurate than those computed at the CCSDTQ level of theory when compared to FCI data.[119] So, the performance of perturbative excitations in coupled cluster theory should be further explored to determine whether CCSD(T) and CCSDT(Q) can be considered as accurate as CCSDT and CCSDTQ, respectively.

In the present work, potential energy surfaces for a subset of homonuclear diatomic molecules are computed at varying levels of theory, and spectroscopic constants are presented. These systems are chosen due to their high symmetry and low cost such that an accurate potential energy surface may be computed with high-order coupled cluster theory. Additionally, this includes C_2 so that these methods may be tested on a system with known multi-reference character. We aim to compare the performance of each coupled cluster method to determine whether perturbative excitations perform better than their full excitation counterparts.

3.2 Methods

Potential energy surfaces for the closed-shell ground states of C_2, N_2, and F_2 were computed using coupled cluster theory with single, double, and perturbative triple excitations [CCSD(T)][85,88,89], with full triple excitations [CCSDT][91–93], with perturbative quadruple excitations [CCSDT(Q)][94,95] and with full quadruple excitations [CCSDTQ][96–99]. For the open-shell triplet ground state of O_2, potential energy surfaces were computed using CCSD(T), CCSDT, ansatz B of CCSDT(Q) derived by Kállay and Gauss [CCSDT(Q)/B][120], and CCSDTQ. Energies, potential energy surfaces, and spectroscopic constants computed with the CCSDT(Q)/B method on the triplet ground state of O_2 may henceforth be referred to as "CCSDT(Q)" for simplicity and consistency in comparison among methods.

Energies using CCSD(T) and CCSDT were computed with a Dunning correlation consistent sextuple-ζ basis set [cc-pV6Z][82]. Energies using CCSDT(Q), CCSDT(Q)/B, and CCSDTQ were computed with a Dunning correlation consistent triple-ζ basis set [cc-pVTZ][82]. All energies were computed through the PSI4 quantum chemistry package and its integration to the MRCC quantum chemistry package when necessary.[102,103,121]

From each potential energy surface for the homonuclear diatomic molecules of interest, four spectroscopic constants were computed at each level of theory: r_e (equilibrium interatomic distance), ω_e (harmonic vibration wavenumber), $\omega_e x_e$ (vibrational anharmonicity constant), and D_e (centrifugal distortion constant). These spectroscopic constants were computed through the PSI4 quantum chemistry package.[121,122]

3.3 Results and Discussion

The computed potential energy surfaces for the ground states of C_2, N_2, O_2 and F_2 are shown in Figures (3.1–3.8). Potential energy surfaces computed at the CCSD(T) and CCSDT levels of theory are shown separately from those computed at the CCSDT(Q) and CCSDTQ levels of theory due to the differences in basis sets used. From inspection, there is almost no difference in the potential energy surfaces of each molecule between CCSD(T) and CCSDT nor between CCSDT(Q) and CCSDTQ.

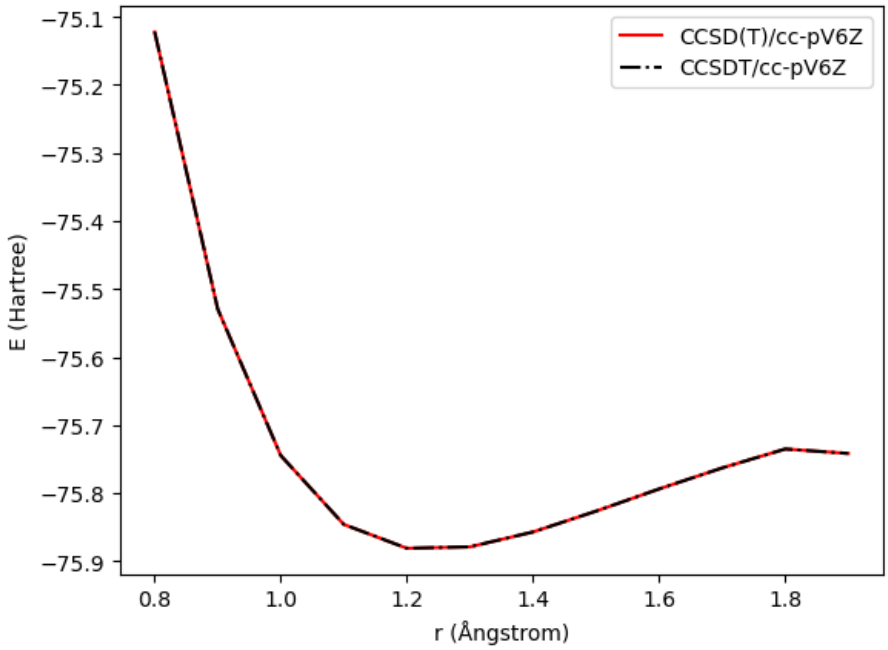

Figure 3.1: Potential energy surface for C_2 at the CCSD(T) and CCSDT levels of theory.

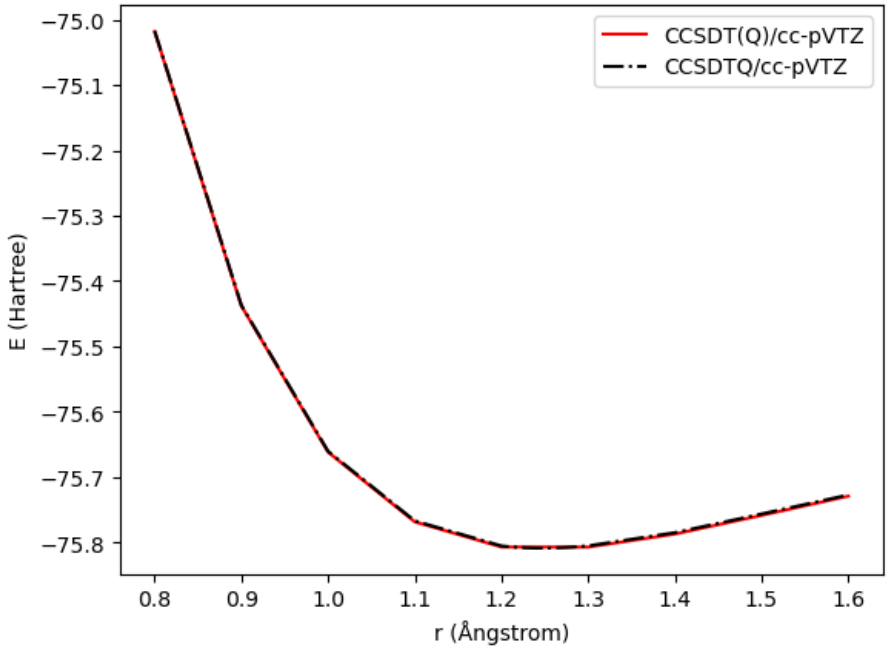

Figure 3.2: Potential energy surface for C_2 at the CCSDT(Q) and CCSDT(Q) levels of theory.

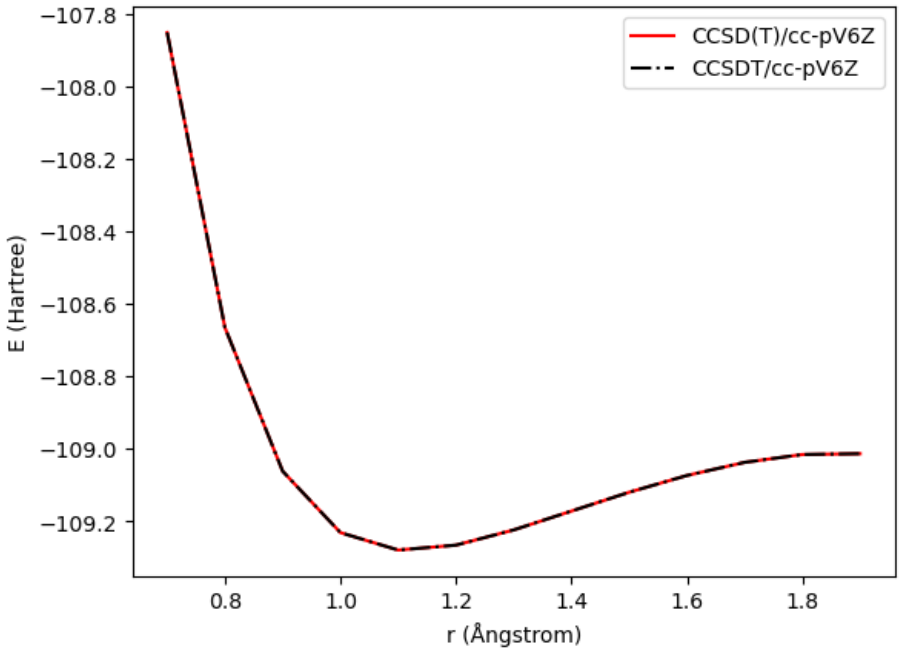

Figure 3.3: Potential energy surface for N_2 at the CCSD(T) and CCSDT levels of theory.

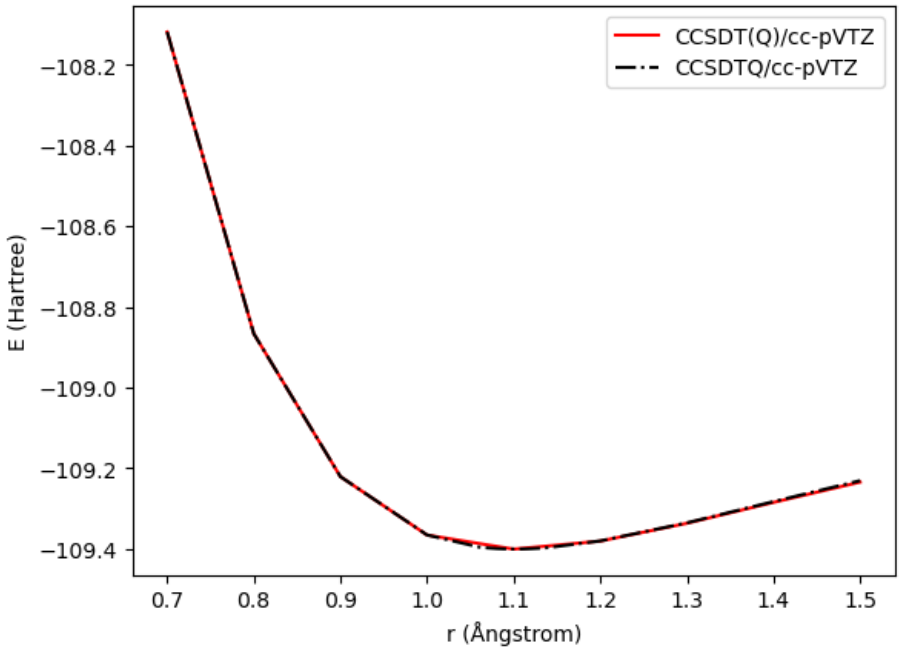

Figure 3.4: Potential energy surface for N_2 at the CCSDT(Q) and CCSDT(Q) levels of theory.

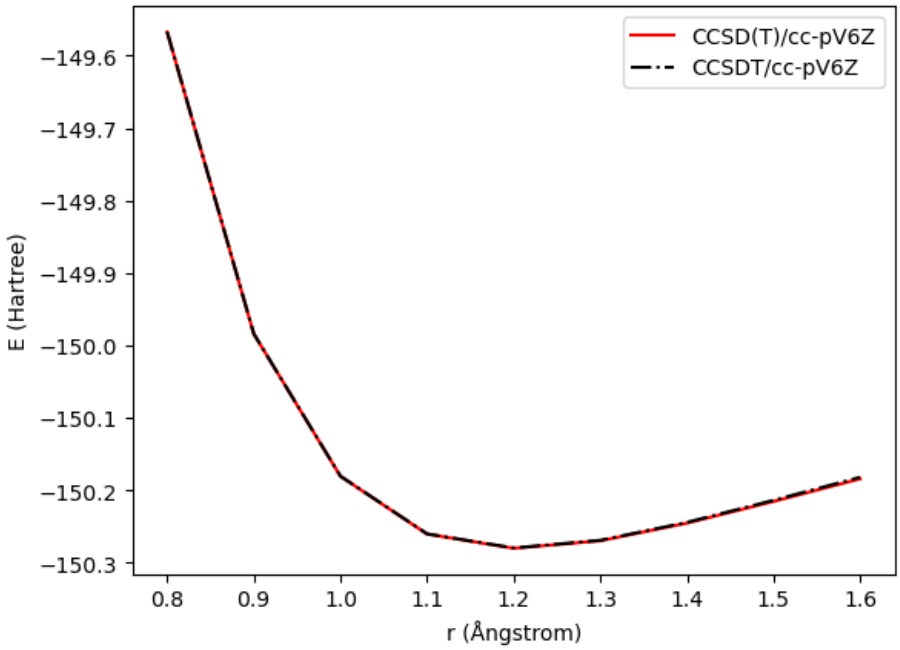

Figure 3.5: Potential energy surface for O_2 at the CCSD(T) and CCSDT levels of theory.

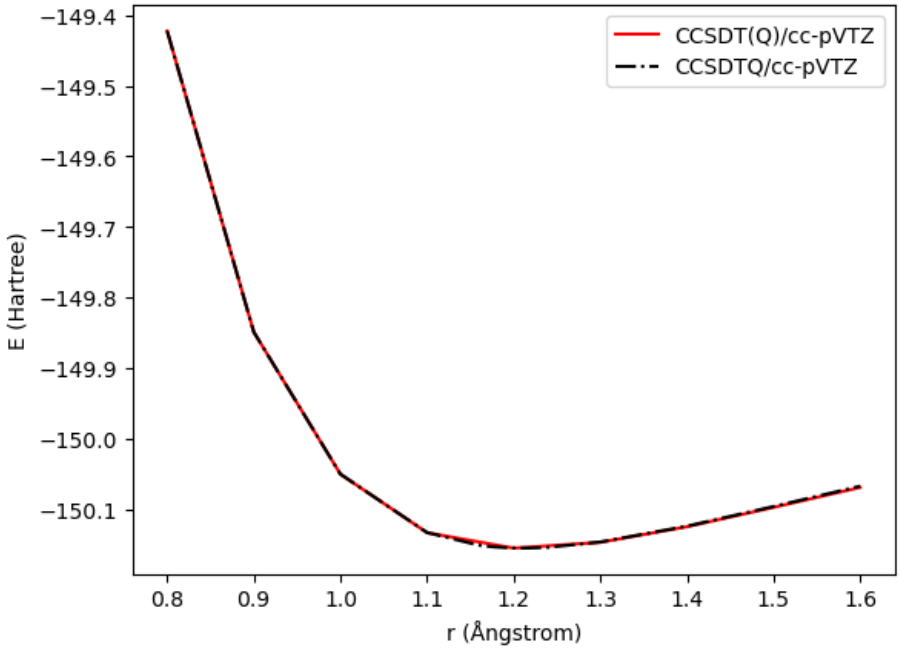

Figure 3.6: Potential energy surface for O_2 at the CCSDT(Q) and CCSDT(Q) levels of theory.

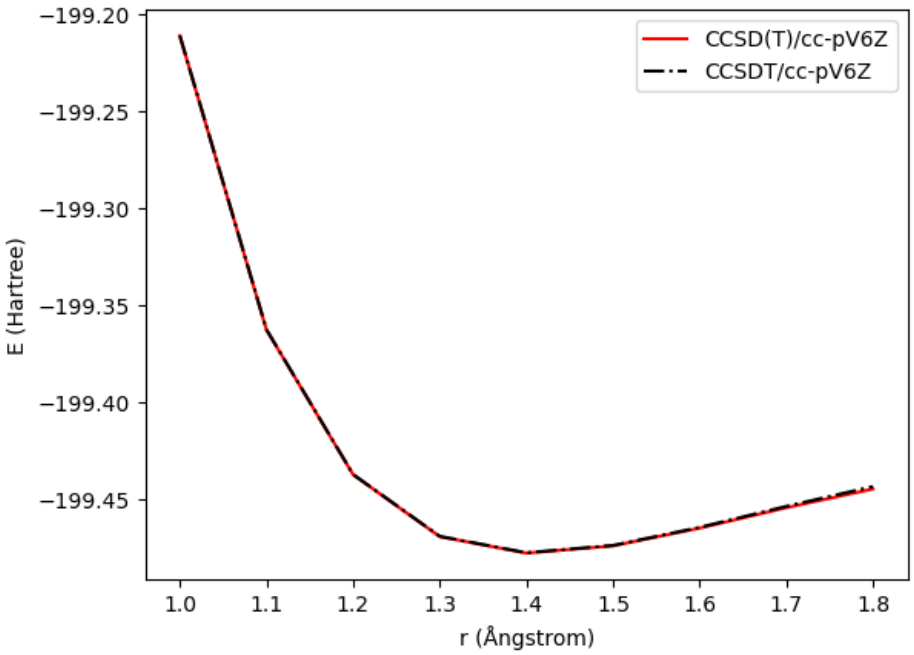

Figure 3.7: Potential energy surface for F_2 at the CCSD(T) and CCSDT levels of theory.

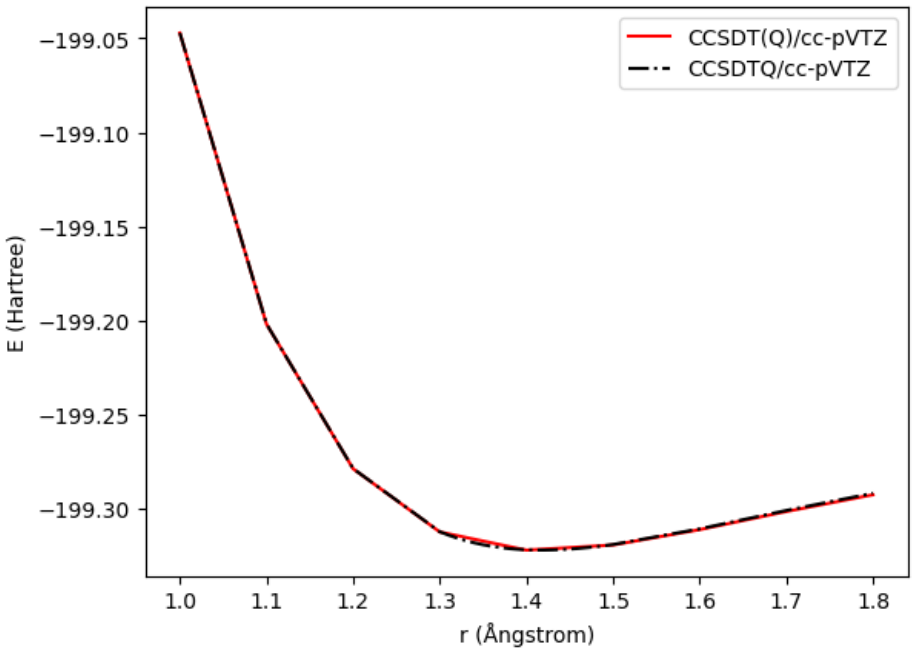

Figure 3.8: Potential energy surface for F_2 at the CCSDT(Q) and CCSDT(Q) levels of theory.

The obtained values for r_e, ω_e , $\omega_e x_e$, and D_e at each level of theory are presented in Tables (3.1–3.4), along with the respective experimental values from the CRC Handbook of Chemistry and Physics[123]. The errors of the obtained constants compared to the literature values are presented in Tables (3.5–3.8).

Table 3.1: Computed r_e values for C_2, N_2, O_2 and F_2 and the literature experimental value. Constants are reported in $\mathring{A}$.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ	Experimental
C_2	1.2383	1.2380	1.2465	1.2457	1.24244
N_2	1.0923	1.0917	1.0969	1.0965	1.09769
O_2	1.1985	1.1978	1.2060	1.2053	1.20752
F_2	1.4045	1.4041	1.4143	1.4135	1.41264

Table 3.2: Computed ω_e values for C_2, N_2, O_2 and F_2 and the literature experimental value. Constants are reported in cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ	Experimental
C_2	1833.89	1836.34	1812.37	1817.57	1855.01
N_2	2405.00	2415.39	2374.09	2380.79	2358.56
O_2	1627.42	1635.42	1571.78	1579.44	1580.19
F_2	933.13	936.69	899.34	903.56	916.93

Table 3.3: Computed $\omega_e x_e$ values for C_2, N_2, O_2 and F_2 and the literature experimental value. Constants are reported in cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ	Experimental
C_2	19.46	19.76	20.60	20.81	13.56
N_2	28.36	27.95	29.99	29.52	14.32
O_2	26.13	25.72	28.64	28.21	11.98
F_2	19.59	19.08	21.00	20.68	11.32

Table 3.4: Computed D_e values for C_2, N_2, O_2 and F_2 and the literature experimental value. Constants are reported in 10^{-6} cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ	Experimental
C_2	7.3	7.3	7.2	7.2	6.96
N_2	5.7	5.7	5.7	5.7	5.737
O_2	4.8	4.7	4.9	4.9	4.839
F_2	3.3	3.3	3.5	3.4	3.3

Table 3.5: Errors of computed r_e values for C_2, N_2, O_2 and F_2 relative to the literature experimental value. Errors are reported in $\mathring{A}$.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ
C_2	–0.0041	–0.0044	0.0041	0.0033
N_2	–0.0054	–0.0060	–0.0011	–0.0015
O_2	–0.0090	–0.0097	–0.0015	–0.0022
F_2	–0.0085	–0.0089	0.0013	0.0005

Table 3.6: Errors of computed ω_e values for C_2, N_2, O_2 and F_2 relative to the literature experimental value. Constants are reported in cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ
C_2	–21.12	–18.67	–42.64	–37.44
N_2	46.44	56.83	15.53	22.23
O_2	47.23	55.23	–8.41	–0.75
F_2	16.20	19.76	–17.59	–13.37

Table 3.7: Errors of computed $\omega_e x_e$ values for C_2, N_2, O_2 and F_2 relative to the literature experimental value. Constants are reported in cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ
C_2	5.90	6.20	7.04	7.25
N_2	14.04	13.63	15.67	15.20
O_2	14.15	13.74	16.66	16.23
F_2	8.27	7.76	9.68	9.36

Table 3.8: Errors of computed D_e values for C_2, N_2, O_2 and F_2 relative to the literature experimental value. Constants are reported in 10^{-6} cm^{-1}.

Molecule	CCSD(T)	CCSDT	CCSDT(Q)	CCSDTQ
C_2	0.3	0.3	0.2	0.2
N_2	0.0	0.0	0.0	0.0
O_2	0.0	–0.1	0.1	0.1
F_2	0.0	0.0	0.2	0.1

3.3.1 Equilibrium Interatomic Distance (r_e)

For r_e, CCSD(T) reported closer values to the experimental values than CCSDT for all molecules. However, CCSDT(Q) values were closer to the experimental values compared to those from CCSDTQ for N_2 and O_2, but were further from the experimental values

for C_2 and F_2. So, methods with perturbative excitations were slightly more accurate in computing r_e values than those without except for CCSDT(Q) on F_2 and the multi-reference C_2. However, due to the minor differences in distances between those computed with and without perturbative excitations, it can be seen that perturbative excitations do accurate approximate r_e values computed with full coupled cluster.

3.3.2 Harmonic Vibration Wavenumber (ω_e)

For ω_e, CCSD(T) reported closer values to the experimental values than CCSDT for all molecules except for C_2, and the computed CCSDT(Q) values were further from the experimental values compared to those from CCSDTQ for all molecules except for N_2. Therefore, for single-reference molecules, perturbative excitations generally performed well. However, for the multi-reference C_2, both CCSD(T) and CCSDT(Q) were less accurate than CCSDT and CCSDTQ, respectively.

3.3.3 Vibration Anharmonicity Constant ($\omega_e x_e$)

CCSD(T) reported $\omega_e x_e$ values further from the experimental values than CCSDT for all molecules except for C_2. Also, the computed CCSDT(Q) values for $\omega_e x_e$ were further from the experimental values compared to those from CCSDTQ for all molecules except for C_2. So, both CCSD(T) and CCSDT(Q) were not as effective at predicting $\omega_e x_e$ at the CCSDT and CCSDTQ levels of theory. The discrepancy in predictive behavior between C_2 and the rest of the diatomic molecules can likely be attributed to the multi-reference character of C_2, although it must be noted that the $\omega_e x_e$ values computed with CCSD(T) and CCSDT(Q) were still more accurate than those computed with CCSDT and CCSDTQ for C_2.

3.3.4 Centrifugal Distortion Constant (D_e)

For D_e, reported values for at the CCSD(T) and CCSDT levels of theory were identical for all molecules besides O_2, for which only CCSD(T) reported the experimental value exactly. Reported values for CCSDT(Q) and CCSDTQ were identical for all molecules besides F_2, for which CCSDTQ reported a closer value to the experimental value. Further analysis on values for D_e is not possible due to the numerical cutoffs for displayed data in the most recent implementation of the software used. However, it can be seen that the methods with perturbative excitations perform well in predicting values using full excitation methods when computing D_e.

3.4 Conclusion

In general, the spectroscopic constants computed for this subset of diatomic molecules at the CCSD(T) level of theory are nearly as accurate, if not more accurate, than those computed at CCSDT. Additionally, the spectroscopic constants computed at the CCSDT(Q) level of theory were nearly as accurate as those computed at CCSDTQ. Because methods with perturbative excitations are more computationally cost-effective methods than their full excitation counterparts, but have shown here to computed spectroscopic constants with similar performance, we suggest that CCSD(T) and CCSDT(Q) are good approximations to CCSDT and CCSDTQ, respectively.

However, due to the discrepancies in behavior between the multi-reference C_2 and the more single-reference diatomic molecules of interest, we suggest that the use of perturbative excitations for multi-reference systems may not be as reliable as their full excitation counterparts of the same order. Therefore, one must use caution when utilizing such methods on multi-reference systems, such as C_2.

It must be noted that the potential energy surfaces, and therefore the spectroscopic constants from them, were not able to be computed at the full configuration interaction limit with a large enough basis set (e.g. cc-pVTZ) to be reported with confidence for any of the molecules of interest. With improved hardware, this may be realized, and further studies may be conducted with FCI computations included to which data computed at each level of theory may be compared.

Chapter 4

Concluding Remarks

In Chapter 1, various quantum chemical methods were discussed, including perturbation theory and how it could be applied to coupled cluster theory. In Chapters 2 and 3, the use of perturbative excitations within coupled cluster theory has been benchmarked through studies of the enthalpies of formation for simple Criegee intermediates as well as of both spectroscopic constants and of potential energy surfaces for homonuclear diatomic molecules.

In the study of Criegee intermediates in Chapter 2, it was shown that coupled cluster energies computed with perturbative quarduple excitations do not agree well with the coupled cluster energies computed without perturbative excitations of the same order for those systems, which can be attributed to the mid multi-reference character of Criegee intermediates. CCSDT(Q) energies over-estimated the contribution of CCSDTQ by a factor of two for energies computed on all four studied Criegee intermediates, with and without the inclusion of optimized virtual orbitals as a cost-saving technique. Because CCSDTQ energies were computed on these systems, highly-accurate enthalpies of formation were still reported for all four Criegee intermediates of interest.

This idea is supported by the work in Chapter 3 on diatomic molecules, in which spectroscopic constants of the multi-reference C_2 computed with perturbative excitations in coupled cluster were often less accurate to experimental values when compared to those computed using coupled cluster methods of the same order but with standard excitation procedures used. However, on molecules with more single-reference character, constants computed with perturbative excitations were often just as accurate, if not more accurate, than their standard coupled cluster counterparts. In cases where CCSD(T) or CCSDT(Q) were more accurate than CCSDT or CCSDTQ respectively for computations on the multi-reference C_2, the opposite was true on the other systems. So, methods with perturbative excitations should be used with caution on systems with discernible multi-reference character.

From these studies and discussions, it has been shown that the use perturbative excitations in coupled cluster theory is may not be as an effective approximation when used on

systems with multi-reference character compared to those with more single-reference character. However, when used on single-reference systems, it is often an accurate, cost-effective approximation to standard coupled cluster theory.

Bibliography

(1) A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Macmillan Publishing Co., Inc., 1982.

(2) I. N. Levine, *Quantum Chemistry, 6th ed.* Prentice Hall, 2008.

(3) R. J. Bartlett, *Ann. Rev. Phys. Chem.*, 1981, **32**, 359–401.

(4) M. L. Leininger, W. D. Allen, H. F. Schaefer III and C. D. Sherrill, *J. Chem. Phys.*, 2000, **112**, 9213–9222.

(5) I. Shavitt and R. J. Bartlett, Cambridge University Press, 2009.

(6) T. Helgaker, P. Horgensen and J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons, 2014.

(7) A. L. East and W. D. Allen, *J. Chem. Phys.*, 1993, **99**, 4638.

(8) N. L. Allinger, J. T. Fermann, W. D. Allen and H. F. Schaefer III, *J. Chem. Phys.*, 1997, **106**, 5143–5150.

(9) A. G. Csaszar, W. D. Allen and H. F. Schaefer III, *J. Chem. Phys.*, 1998, **108**, 9751.

(10) J. M. Gonzales, C. Pak, R. S. Cox, W. D. Allen, H. F. Schaefer Iii, A. G. Császár and G. Tarczay, *Chem. Eur. J.*, 2003, **9**, 2173–2192.

(11) A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez and J. F. Stanton, *J. Chem. Phys.*, 2004, **121**, 11599–11613.

(12) A. Karton, E. Rabinovich, J. M. Martin and B. Ruscic, *J. Chem. Phys.*, 2006, **125**, 144108.

(13) D. Feller, *J. Chem. Phys.*, 1992, **96**, 6104.

(14) D. Feller, *J. Chem. Phys.*, 1993, **98**, 7059.

(15) T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639.

(16) P. Neogrády, M. Pitoňák and M. Urban, *Mol. Phys.*, 2005, **103**, 2141.

(17) Z. Rolik and M. Kállay, *J. Chem. Phys.*, 2011, **134**, 124111.

(18) R. Criegee, *Angew. Chem. Int. Ed.*, 1975, **14**, 745.

(19) R. Criegee and G. Wenner, *Liebigs Ann.*, 1949, **564**, 9–15.

(20) M. Khan, C. Percival, R. Caravan, C. Taatjes and D. Shallcross, *Environ. Sci. Process. Impacts*, 2018, **20**, 437.

(21) M. S. Alam, A. R. Rickard, M. Camredon, K. P. Wyche, T. Carr, K. E. Hornsby, P. S. Monks and W. J. Bloss, *J. Phys. Chem. A*, 2013, **117**, 12468.

(22) A. S. Hasson, G. Orzechowska and S. E. Paulson, *J. Geophys. Res. Atmos.*, 2001, **106**, 34131.

(23) J. T. Herron, R. I. Martinez and R. E. Huie, *Int. J. Chem. Kinet.*, 1982, **14**, 201.

(24) O. Horie and G. K. Moortgat, *Atmos. Environ.*, 1991, **25**, 1881.

(25) F. Liu, J. M. Beames and M. I. Lester, *J. Chem. Phys.*, 2014, **141**, 234312.

(26) T. A. Stephenson and M. I. Lester, *Int. Rev. Phys. Chem.*, 2020, **39**, 1–33.

(27) T. L. Nguyen, H. Lee, D. A. Matthews, M. C. McCarthy and J. F. Stanton, *J. Phys. Chem. A*, 2015, **119**, 5524.

(28) W. Chao, J.-T. Hsieh, C.-H. Chang and J. J.-M. Lin, *Science*, 2015, **347**, 751.

(29) M. C. Smith, C.-H. Chang, W. Chao, L.-C. Lin, K. Takahashi, K. A. Boering and J. J.-M. Lin, *J. Phys. Chem. Lett.*, 2015, **6**, 2708.

(30) A. B. Ryzhkov and P. A. Ariya, *J. Chem. Phys. Lett.*, 2006, **419**, 479.

(31) T. R. Lewis, M. A. Blitz, D. E. Heard and P. W. Seakins, *Phys. Chem. Chem. Phys.*, 2015, **17**, 4859.

(32) L. Sheps, B. Rotavera, A. J. Eskola, D. L. Osborn, C. A. Taatjes, K. Au, D. E. Shallcross, M. A. H. Khan and C. J. Percival, *Phys. Chem. Chem. Phys.*, 2017, **19**, 21970.

(33) S. Hatakeyama and H. Akimoto, *Res. Chem. Intermed.*, 1994, **20**, 503.

(34) J. P. Misiewicz, S. N. Elliott, K. B. Moore and H. F. Schaefer, *Phys. Chem. Chem. Phys.*, 2018, **20**, 7479.

(35) Y. Liu, C. Yin, M. C. Smith, S. Liu, M. Chen, X. Zhou, C. Xiao, D. Dai, K. Takahashi, W. Dong and X. Yang, *Phys. Chem. Chem. Phys.*, 2018, **20**, 29669.

(36) R. Chhantyal-Pun, R. J. Shannon, D. P. Tew, R. L. Caravan, M. Duchi, C. Wong, A. Ingham, C. Feldman, M. R. McGillen, M. A. H. Khan, I. O. Antonov, B. Rotavera, K. Ramasesha, D. L. Osborn, C. A. Taatjes, C. J. Percival, D. E. Shallcross and A. J. Orr-Ewing, *Phys. Chem. Chem. Phys.*, 2019, **21**, 14042.

(37) Y. Wei, F. Xu, X. Ma, W. Wang and Q. Zhang, *Chem. Phys. Lett.*, 2020, **743**, 137140.

(38) R. L. Caravan, M. A. H. Khan, B. Rotavera, E. Papajak, I. O. Antonov, M.-W. Chen, K. Au, W. Chao, D. L. Osborn, C. J. Percival, D. E. Shallcross and C. A. Taatjes, *Faraday Discuss.*, 2017, **200**, 313.

(39) B. Ouyang, M. W. McLeod, R. L. Jones and W. J. Bloss, *Phys. Chem. Chem. Phys.*, 2013, **15**, 17070.

(40) H.-L. Huang, W. Chao and J. J.-M. Lin, *Proc. Natl. Acad. Sci.*, 2015, **112**, 10857.

(41) M. Sipilä, T. Jokinen, T. Berndt, S. Richters, R. Makkonen, N. Donahue, R. Mauldin Iii, T. Kurtén, P. Paasonen, N. Sarnela, M. Ehn, H. Junnien, M. P. Rissanen, J. Thornton, F. Stratmann, H. Herrmann, D. R. Worsnop, M. Kulmala, V.-M. Kerminen and Petäjä, *Atmos. Chem. Phys.*, 2014, **14**, 12143.

(42) R. I. Martinez and J. T. Herron, *J. Environ. Sci. Health A*, 1981, **16**, 623.

(43) A. Jalan, J. W. Allen and W. H. Green, *Phys. Chem. Chem. Phys.*, 2013, **15**, 16841.

(44) H. J. Tobias and P. J. Ziemann, *J. Phys. Chem. A*, 2001, **105**, 6129.

(45) N. A. Watson, J. A. Black, T. M. Stonelake, P. J. Knowles and J. M. Beames, *J. Phys. Chem. A*, 2018, **123**, 218.

(46) G. J. Aroeira, A. S. Abbott, S. N. Elliott, J. M. Turney and H. F. Schaefer, *Phys. Chem. Chem. Phys.*, 2019, **21**, 17760.

(47) S. V. Tadayon, E. S. Foreman and C. Murray, *J. Phys. Chem. A*, 2017, **122**, 258.

(48) R. Chhantyal-Pun, B. Rotavera, M. R. McGillen, M. A. H. Khan, A. J. Eskola, R. L. Caravan, L. Blacker, D. P. Tew, D. L. Osborn, C. J. Percival, C. A. Taatjes, D. E. Shallcross and A. J. Orr-Ewing, *ACS Earth Space Chem.*, 2018, **2**, 833.

(49) O. Welz, A. J. Eskola, L. Sheps, B. Rotavera, J. D. Savee, A. M. Scheer, D. L. Osborn, D. Lowe, A. Murray Booth, P. Xiao, M. A. H. Khan, C. J. Percival, D. E. Shallcross and C. A. Taatjes, *Angew. Chem. Int. Ed.*, 2014, **53**, 4547.

(50) S. Zhou, S. Joudan, M. W. Forbes, Z. Zhou and J. P. Abbatt, *Environ. Sci. Technol. Lett.*, 2019, **6**, 243.

(51) D. L. Osborn and C. A. Taatjes, *Int. Rev. Phys. Chem.*, 2015, **34**, 309–360.

(52) T. B. Nguyen, G. S. Tyndall, J. D. Crounse, A. P. Teng, K. H. Bates, R. H. Schwantes, M. M. Coggon, L. Zhang, P. Feiner, D. O. Milller, K. M. Skaog, J. C. Riviera-Rios, M. Dorris, K. F. Olson, A. Koss, R. J. Wild, S. S. Brown, A. H. Goldstein, J. A. de Gouw, W. H. Brune, F. N. Keutsch, J. H. Seinfeld and P. O. Wennberg, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10241.

(53) M. Nakajima and Y. Endo, *J. Chem. Phys.*, 2015, **143**, 164307.

(54) O. Welz, J. D. Savee, D. L. Osborn, S. S. Vasu, C. J. Percival, D. E. Shallcross and C. A. Taatjes, *Science*, 2012, **335**, 204.

(55) C. A. Taatjes, O. Welz, A. J. Eskola, J. D. Savee, A. M. Scheer, D. E. Shallcross, B. Rotavera, E. P. Lee, J. M. Dyke, D. K. Mok, D. L. Osborn and C. J. Percival, *Science*, 2013, **340**, 177.

(56) T. Berndt, T. Jokinen, R. L. Mauldin III, T. Petäjä, H. Herrmann, H. Junninen, P. Paasonen, D. R. Worsnop and M. Sipilä, *J. Phys. Chem. Lett.*, 2012, **3**, 2892.

(57) L. Vereecken, H. Harder and A. Novelli, *Physical Chemistry Chemical Physics*, 2012, **14**, 14682–14695.

(58) C. A. Taatjes, G. Meloni, T. M. Selby, A. J. Trevitt, D. L. Osborn, C. J. Percival and D. E. Shallcross, *J. Am. Chem. Soc.*, 2008, **130**, 11883.

(59) R. Chhantyal-Pun, M. A. H. Khan, C. A. Taatjes, C. J. Percival, A. J. Orr-Ewing and D. E. Shallcross, *Int. Rev. Phys. Chem.*, 2020, **39**, 385–424.

(60) C. A. Taatjes, *Annu. Rev. Phys. Chem.*, 2017, **68**, 183–207.

(61) M. C. Smith, W.-L. Ting, C.-H. Chang, K. Takahashi, K. A. Boering and J. J.-M. Lin, *J. Chem. Phys.*, 2014, **141**, 074302.

(62) L. Sheps, A. M. Scully and K. Au, *Phys. Chem. Chem. Phys.*, 2014, **16**, 26701–26706.

(63) J. M. Beames, F. Liu, L. Lu and M. I. Lester, *J. Chem. Phys.*, 2013, **138**, 20045.

(64) R. Chhantyal-Pun, O. Welz, J. D. Savee, A. J. Eskola, E. P. Lee, L. Blacker, H. R. Hill, M. Ashcroft, M. A. H. Khan, G. C. Lloyd-Jones et al., *The Journal of Physical Chemistry A*, 2017, **121**, 4–15.

(65) F. Liu, J. M. Beames, A. M. Green and M. I. Lester, *The Journal of Physical Chemistry A*, 2014, **118**, 2298–2306.

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

(67) H.-L. Huang, W. Chao and J. J.-M. Lin, *Proc. Natl. Acad. Sci.*, 2015, **112**, 10857.

(68) D. A. Matthews, *J. Chem. Theory Comput.*, 2020, **16**, 6195.

(69) M. T. Nguyen, T. L. Nguyen, V. T. Ngan and H. M. T. Nguyen, *Chem. Phys. Lett.*, 2007, **448**, 183.

(70) L. B. Harding and W. A. Goddard, *J. Am. Chem. Soc.*, 1978, **100**, 7180.

(71) M. Olzmann, E. Kraka, D. Cremer, R. Gutbrod and S. Andersson, *J. Phys. Chem. A*, 1997, **101**, 9421.

(72) D. Cremer, J. Gauss, E. Kraka, J. F. Stanton and R. J. Bartlett, *Chem. Phys. Lett.*, 1993, **209**, 547.

(73) M. Kettner, A. Karton, A. McKinley and D. Wild, *Chem. Phys. Lett.*, 2015, **621**, 193.

(74) A. Karton and J. M. L. Martin, *J. Chem. Phys.*, 2012, **136**, 124114.

(75) `https://atct.anl.gov`, 2020.

(76) CFOUR, a quantum chemical program package written by J.F. Stanton, J. Gauss, L. Cheng, M.E. Harding, D.A. Matthews, P.G. Szalay with contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke, Y.J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. Kopper, W.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, T. Nottoli, D.P. O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

(77) J. Almlöf and P. R. Taylor, *J. Chem. Phys.*, 1987, **86**, 4070.

(78) J. Almlöf and P. R. Taylor, *J. Chem. Phys.*, 1990, **92**, 551.

(79) M. S. Schuurman, S. R. Muir, W. D. Allen and H. F. Schaefer III, *J. Chem. Phys.*, 2004, **120**, 11586.

(80) D. R. Hartree, *Proc. Camb. Philos. Soc.*, 1928, **24**, 89.

(81) V. Fock, *Z. Phys.*, 1930, **61**, 126.

(82) T. H. Dunning Jr, *J. Chem. Phys.*, 1989, **90**, 1007.

(83) R. J. Bartlett and D. M. Silver, *Phys. Rev. A*, 1974, **10**, 1927.

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

(85) G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice and H. F. Schaefer III, *J. Chem. Phys.*, 1987, **86**, 2881.

(86) J. F. Stanton, J. Gauss, J. D. Watts and R. J. Bartlett, *J. Chem. Phys.*, 1991, **94**, 4334–4345.

(87) C. Hampel, K. A. Peterson and H.-J. Werner, *Chem. Phys. Lett.*, 1992, **190**, 1–12.

(88) K. Raghavachari, *Chem. Phys. Lett*, 2013, **589**, 35–36.

(89) R. J. Bartlett, J. Watts, S. Kucharski and J. Noga, *Chem. Phys. Lett*, 1990, **165**, 513.

(90) J. F. Stanton, *Chem. Phys. Lett.*, 1997, **281**, 130–134.

(91) J. Noga and R. J. Bartlett, *J. Chem. Phys.*, 1987, **86**, 7041.

(92) J. Noga and R. Bartlett, *J. Chem. Phys.*, 1988, **89**, 3401.

(93) G. E. Scuseria and H. F. Schaefer III, *Chem. Phys. Lett*, 1988, **152**, 382.

(94) Y. J. Bomble, J. F. Stanton, M. Kállay and J. Gauss, *J. Chem. Phys.*, 2005, **123**, 054101.

(95) M. Kállay and J. Gauss, *J. Chem. Phys.*, 2005, **123**, 214105.

(96) S. A. Kucharski and R. J. Bartlett, *Theor. Chem. Acc.*, 1991, **80**, 387.

(97) N. Oliphant and L. Adamowicz, *J. Chem. Phys.*, 1991, **95**, 6645.

(98) S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.*, 1992, **97**, 4282.

(99) D. A. Matthews and J. F. Stanton, *J. Chem. Phys.*, 2015, **142**, 064108.

(100) M. Kállay and P. R. Surján, *J. Chem. Phys.*, 2001, **115**, 2945.

(101) J. Olsen, *J. Chem. Phys.*, 2000, **113**, 7140.

(102) M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Segedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei and Á. Ganyecz, *J. Chem. Phys.*, 2020, **152**, 074107.

(103) MRCC, a quantum chemical program suite written by M. Kállay, P. R. Nagy, D. Mester, L. Gyevi-Nagy, J. Csóka, P. B. Szabó, Z. Rolik, G. Samu, J. Csontos, B. Hégely, Á. Ganyecz, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and R. A. Horváth. See www.mrcc.hu.

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

(105) K. G. Dyall, *J. Chem. Phys.*, 1997, **106**, 9618.

(106) W. Kutzelnigg and W. Liu, *J. Chem. Phys.*, 2005, **123**, 241102.

(107) W. Liu and D. Peng, *J. Chem. Phys.*, 2006, **125**, 044102.

(108) M. Iliaš and T. Saue, *J. Chem. Phys.*, 2007, **126**, 064102.

(109) J. Sikkema, L. Visscher, T. Saue and M. Iliaš, *J. Chem. Phys.*, 2009, **131**, 124116.

(110) W. Liu and D. Peng, *J. Chem. Phys.*, 2009, **131**, 031104.

(111) J. Čížek, *J. Chem. Phys.*, 1966, **45**, 4256–4266.

(112) J. Čížek, *Adv. Chem Phys.*, 1969, **14**, 35–89.

(113) M. Kodrycka and K. Patkowski, *J. Chem. Phys.*, 2019, **151**, 070901.

(114) M. Lesiuk, *J. Chem. Theory Comput.*, 2022, **18**, 6537–6556.

(115) J. R. Lane, *J. Chem. Theory Comput.*, 2013, **9**, 316–323.

(116) J. Sordo, *J. Chem. Phys.*, 2001, **114**, 1974–1980.

(117) P. A. Denis, *Chem. Phys. Lett.*, 2004, **395**, 12–20.

(118) J. M. Begley, G. J. R. Aroeira, J. M. Turney, G. E. Douberly and H. F. Schaefer III, *J. Chem. Phys.*, 2023, **158**, 034302.

(119) L. Simova, J. Rezac and P. Hobza, *J. Chem. Theory Comput.*, 2013, **9**, 3420–3428.

(120) M. Kállay and J. Gauss, *J. Chem. Phys.*, 2008, **129**, 144101.

(121) D. G. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer III, A. Y. Sokolov, K. Patkowski, A. E. DePrince III, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Phys.*, 2020, **152**, 184108.

(122) J. D. Bender, S. Doraiswamy, D. G. Truhlar and G. V. Candler, *J. Chem. Phys.*, 2014, **140**, 054302.

(123) J. R. Rumble, *Handbook of Chemistry and Physics, 106th ed.* CRC Press, Taylor & Francis Group, 2025.