

COUPLED-CLUSTER AND EQUATION-OF-MOTION COUPLED-CLUSTER  
THEORIES: APPLICATIONS TO PHOTOCHEMISTRY AND CATALYSIS AND  
ALGORITHMIC ADVANCES

By

Jared A. Hansen

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Chemistry - Doctor of Philosophy

2015

## ABSTRACT

# COUPLED-CLUSTER AND EQUATION-OF-MOTION COUPLED-CLUSTER THEORIES: APPLICATIONS TO PHOTOCHEMISTRY AND CATALYSIS AND ALGORITHMIC ADVANCES

By

Jared A. Hansen

Understanding electronic excitation, photoelectron, and multiphoton ionization spectra, particularly those involving dark and strongly correlated states, or transition metals, as well as catalytic, structural, and electronic properties of gold nanoparticles pose significant challenges for theory and experiment. The existing experimental techniques may not be sufficiently powerful to provide definitive information on their own, whereas an accurate treatment of the relevant many-electron correlation effects required in theoretical analyses may be far from obvious. In this dissertation, we describe several high-level *ab initio* computational studies employing the completely renormalized (CR) and active-space coupled-cluster (CC) and equation-of-motion CC (EOMCC) approaches and the extensions of the EOMCC theory to open-shell systems around closed shells defining the electron-attached (EA) and ionized (IP) EOMCC frameworks, which demonstrate the transformative role these novel electronic structure methods developed in our group have played in understanding the previously unexplained experiments and phenomena. They include (i) challenging electronic spectra of the CNC, C<sub>2</sub>N, N<sub>3</sub>, and NCO molecules and the photoelectron spectrum of Au<sub>3</sub><sup>-</sup> nanoparticle examined with the EA- and IP-EOMCC approaches, especially those invented in our group, (ii) the discovery of the doubly excited state of azulene below the ionization threshold, which mediates the 1 + 2' multiphoton ionization experiments resulting in clear Rydberg finger-

print spectra, where the CR-EOMCC formalism developed in our group played a crucial role, (iii) the detailed investigation of the mechanism and energetics of the aerobic oxidation of methanol on  $\text{Au}_8^-$  particle, which benefited from the application of the ground-state CR-CC methodology, developed by our group as well, and (iv) definitive CR-CC and active-space CC studies showing that the ground state of 1,2,3,4-cyclobutanetetraone, which is characterized by densely spaced low-lying states, is a triplet, in agreement with the recently recorded photodetachment spectrum. These cutting-edge computational studies are accompanied by advances in CC/EOMCC algorithms and methodologies, including the development of parallel numerical energy gradients and second derivatives for fast geometry optimizations and harmonic vibrational frequency calculations at any CC/EOMCC level, allowing us to establish the geometries and relative energies of the low-energy isomers of the controversial  $\text{Au}_8$  particle, and the implementation of the unrestricted Hartree-Fock-based (UHF-based) CR-CC(2,3) approach, allowing us to show that unlike the popular CCSD(T) approach, which is very sensitive to the type of the reference determinant employed in the calculations, failing in bond-breaking situations when the restricted Hartree-Fock (RHF) reference is used and displaying poor behavior at intermediate nuclear separations with UHF references, its CR-CC(2,3) counterpart provides a robust description regardless of the reference type (RHF or UHF), with the spin-adapted RHF-based CR-CC(2,3) results being most accurate in the examined bond dissociation cases.

This is dedicated to my lovely wife Pam and daughter Kaitlyn.

## ACKNOWLEDGMENTS

I would like to thank my PhD. advisor, Professor Piotr Piecuch, for his patience over many long hours of intense debate and discussions as he taught and guided me. I have learned what it means and will take to be a better scientist thanks to his guiding hand and example. I am forever grateful for his continuous support and everything he has done for me.

I would like to thank the other members of my committee, Professor James F. Harrison, Professor James E. Jackson, and Professor Benjamin G. Levine for their advice, suggestions, support, and patience in overseeing my graduate studies. Particularly, I would like to thank Professor Benjamin G. Levine for helping me to contribute to the CIOpt code and for many useful discussions we had that allowed me to better understand technical issues related to parallel numerical derivatives. I also owe a debt of gratitude to Professor Masahiro Ehara, who, through Professor Piotr Piecuch, invited me to spend a few months at the Institute for Molecular Science in Okazaki, Japan, as well as for several stimulating discussions that have resulted in improvements of some of the results presented in this thesis.

To the current and past members of the Piecuch research group, Dr. Jesse J. Lutz, Dr. Wei Li, Dr. Jun Shen, Mr. Nicholas P. Bauman, Mr. Adeayo O. Ajala, and Mr. Jorge Emiliano Deustua, for all the help they have provided me over the years and the many discussions we have had about scientific matters.

# TABLE OF CONTENTS

|                                                                                                                                                                                                   |             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>LIST OF TABLES</b>                                                                                                                                                                             | <b>viii</b> |
| <b>LIST OF FIGURES</b>                                                                                                                                                                            | <b>ii</b>   |
| <b>Chapter 1 Introduction</b>                                                                                                                                                                     | <b>1</b>    |
| <b>Chapter 2 Project objectives</b>                                                                                                                                                               | <b>10</b>   |
| <b>Chapter 3 Applications of full and active-space electron attached and ionized equation-of-motion coupled-cluster methods</b>                                                                   | <b>11</b>   |
| 3.1 Theory                                                                                                                                                                                        | 12          |
| 3.2 Geometries and adiabatic excitation energies of the low-lying valence states of CNC, C <sub>2</sub> N, N <sub>3</sub> , and NCO                                                               | 19          |
| 3.2.1 Computational details                                                                                                                                                                       | 19          |
| 3.2.2 Results                                                                                                                                                                                     | 24          |
| 3.2.2.1 EA-EOMCC results for CNC and C <sub>2</sub> N                                                                                                                                             | 25          |
| 3.2.2.2 IP-EOMCC results for NCO and N <sub>3</sub>                                                                                                                                               | 42          |
| 3.3 Coupled-cluster interpretation of the photoelectron spectrum of Au <sub>3</sub> <sup>-</sup>                                                                                                  | 55          |
| 3.3.1 Introductory remarks                                                                                                                                                                        | 55          |
| 3.3.2 IP-EOMCC extrapolation scheme                                                                                                                                                               | 57          |
| 3.3.3 Results                                                                                                                                                                                     | 64          |
| <b>Chapter 4 Applications of completely renormalized coupled-cluster and equation-of-motion coupled-cluster approaches</b>                                                                        | <b>75</b>   |
| 4.1 Theory                                                                                                                                                                                        | 75          |
| 4.2 Discovery of the doubly excited state that mediates the photoionization of azulene                                                                                                            | 83          |
| 4.2.1 Background information and performed calculations                                                                                                                                           | 83          |
| 4.2.2 Results                                                                                                                                                                                     | 91          |
| 4.3 Aerobic oxidation of methanol to formic acid on Au <sub>8</sub> <sup>-</sup> : Benchmark analysis based on completely renormalized coupled-cluster and density functional theory calculations | 100         |
| 4.3.1 Background information and scope of the work                                                                                                                                                | 100         |
| 4.3.2 Molecular model                                                                                                                                                                             | 105         |
| 4.3.3 Electronic structure methods                                                                                                                                                                | 108         |
| 4.3.4 Results and discussion                                                                                                                                                                      | 114         |
| 4.3.4.1 Coupled-cluster calculations                                                                                                                                                              | 124         |

|                     |                                                                                                                                                     |            |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 4.3.4.2             | Benchmarking DFT against the reference CR-CC(2,3),D data                                                                                            | 129        |
| 4.4                 | Coupled-Cluster and Multireference Configuration Interaction Studies of the<br>Low-Lying Electronic States of 1,2,3,4-Cyclobutanetetraone . . . . . | 136        |
| 4.4.1               | Background information and scope of the work . . . . .                                                                                              | 136        |
| 4.4.2               | Computational details . . . . .                                                                                                                     | 145        |
| 4.4.3               | Results and discussion . . . . .                                                                                                                    | 153        |
| <b>Chapter 5</b>    | <b>Algorithmic advances . . . . .</b>                                                                                                               | <b>165</b> |
| 5.1                 | Parallel numerical derivatives . . . . .                                                                                                            | 165        |
| 5.1.1               | Gradients for geometry optimizations . . . . .                                                                                                      | 168        |
| 5.1.2               | Hessians for harmonic vibrational analyses . . . . .                                                                                                | 176        |
| 5.1.3               | Application to the low-energy isomers of Au <sub>8</sub> . . . . .                                                                                  | 180        |
| 5.2                 | Unrestricted implementation of CR-CC(2,3) . . . . .                                                                                                 | 191        |
| 5.2.1               | Algorithm . . . . .                                                                                                                                 | 193        |
| 5.2.2               | Numerical examples . . . . .                                                                                                                        | 195        |
| 5.2.2.1             | The HF molecule . . . . .                                                                                                                           | 195        |
| 5.2.2.2             | The F <sub>2</sub> molecule . . . . .                                                                                                               | 197        |
| 5.2.2.3             | The H <sub>2</sub> O <sub>2</sub> molecule . . . . .                                                                                                | 205        |
| 5.2.2.4             | The C <sub>2</sub> H <sub>6</sub> molecule . . . . .                                                                                                | 208        |
| <b>Chapter 6</b>    | <b>Conclusions and future outlook . . . . .</b>                                                                                                     | <b>215</b> |
| <b>BIBLIOGRAPHY</b> | <b>. . . . .</b>                                                                                                                                    | <b>219</b> |

# LIST OF TABLES

|            |                                                                                                                                                                                                                                                                    |    |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 3.1: | Total and adiabatic excitation energies for the ground and low-lying excited states of CNC, as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ ( $x = D, T, Q$ ) basis sets and extrapolating to the CBS limit. . . . .         | 26 |
| Table 3.2: | Total and adiabatic excitation energies for the ground and low-lying excited states of C <sub>2</sub> N obtained with the various EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ ( $x = D, T, Q$ ) basis sets and extrapolating to the CBS limit. . . . .  | 33 |
| Table 3.3: | Comparison of the optimized equilibrium geometries for the low-lying states of CNC and C <sub>2</sub> N, as obtained with the EA-EOMCC and SAC-CI-SDT- <i>R</i> /PS approaches using the DZP[4s2p1d] and cc-pVxZ ( $x = D, T, Q$ ) basis sets. . . . .             | 39 |
| Table 3.4: | Total and adiabatic excitation energies for the ground and low-lying excited states of NCO, as obtained with the different IP-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ ( $x=D,T,Q$ ) basis sets and extrapolating to the CBS limit. . . . .             | 44 |
| Table 3.5: | Total and adiabatic excitation energies for the ground and low-lying excited states of N <sub>3</sub> , as obtained with the different IP-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ ( $x=D,T,Q$ ) basis sets and extrapolating to the CBS limit. . . . . | 48 |
| Table 3.6: | Comparison of the optimized equilibrium geometries for the low-lying states of N <sub>3</sub> and NCO, as obtained with the IP-EOMCC and SAC-CI-SDT- <i>R</i> /PS approaches using the DZP [4s2p1d] and cc-pVxZ ( $x=D,T,Q$ ) basis sets. . . . .                  | 53 |
| Table 3.7: | Scalar relativistic ionization energies (in eV) of Au from Au <sup>-</sup> computed using the IP-EOMCC approaches; comparison with experiment and experimentally derived estimates. . . . .                                                                        | 59 |



|            |                                                                                                                                                                                                                                                                                                                             |     |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 3.8: | Vertical <sup>a</sup> (V) and adiabatic <sup>b</sup> (Ad) IEs (in eV) of $\text{Au}_3^-$ with respect to its ground state computed using the scalar relativistic IP-EOMCCSD( $2h-1p$ )/aug-cc-pV $x$ Z-PP(+CV) ( $x = \text{D, T}$ ) and IP-EOMCCSD( $3h-2p$ )/aug-cc-pVDZ-PP approaches.                                   | 66  |
| Table 4.1: | Bond lengths (in Å) defining the calculated and experimental geometries of azulene in its ground electronic state $S_0$ ( $X^1A_1$ ) (for the meaning of various distances, see Fig. 4.1).                                                                                                                                  | 95  |
| Table 4.2: | Vertical excitation energies (in eV) and REL values for excited states of azulene found in this work.                                                                                                                                                                                                                       | 96  |
| Table 4.3: | The absolute values of the reference ( $r_{\mu,0}$ ) and leading singly excited ( $r_a^i$ ) and doubly excited ( $r_{ab}^{ij}$ ) amplitudes, along with the corresponding orbital information, defining the EOMCCSD wave functions of various excited states of azulene obtained using the 6-31G(d) and cc-pVDZ basis sets. | 97  |
| Table 4.4: | Oscillator strengths for the various transitions in azulene obtained from the EOMCCSD calculations using the 6-31G(d) and cc-pVDZ basis sets and the available CASSCF(10,10)/6-31G(d) [262] and experimental [273] results.                                                                                                 | 98  |
| Table 4.5: | Energy differences (in kcal/mol) for the various species along pathway I computed using selected CC and DFT approaches and the BS1 basis set. <sup>a</sup>                                                                                                                                                                  | 118 |
| Table 4.6: | Energy differences (in kcal/mol) for the various species along pathway I computed using selected CC and DFT approaches and the BS2 basis set. <sup>a</sup>                                                                                                                                                                  | 119 |
| Table 4.7: | Energy differences (in kcal/mol) for the various species along pathway I computed at the CCSD and CR-CC(2,3) levels using the BS1 basis set and two different schemes for obtaining the canonical ROHF orbitals.                                                                                                            | 121 |
| Table 4.8: | Comparisons of energy differences (in kcal/mol) for the various species along pathway I obtained using two different ways of extrapolating the CC/BS2 energetics, <sup>a</sup> represented by Eqs. (4.28) and (4.29).                                                                                                       | 123 |
| Table 4.9: | Energies of the four lowest-energy electronic states of $\text{C}_4\text{O}_4$ at the UB3LYP/6-31G(d), CASSCF(2,2)/6-31G(d), and CASSCF(16,16)/6-31G(d) geometries obtained in Ref. [327]. <sup>a</sup>                                                                                                                     | 148 |

|             |                                                                                                                                                                                                                                                                                                                                                              |     |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 4.10: | Orbital structure of CASSCF(16,16) and MRCI(Q)(14,13) active spaces.<br>151                                                                                                                                                                                                                                                                                  |     |
| Table 4.11: | Bond lengths of optimized C <sub>4</sub> O <sub>4</sub> <i>D</i> <sub>4h</sub> geometries (in Å). <sup>a</sup> . . . . .                                                                                                                                                                                                                                     | 152 |
| Table 4.12: | Energies of the low-lying electronic states of C <sub>4</sub> O <sub>4</sub> at the geometries<br>obtained with CC methods. <sup>a</sup> . . . . .                                                                                                                                                                                                           | 156 |
| Table 5.1:  | Wall time (hours) comparison of parallel and serial numerical geom-<br>etry optimizations of several molecules. . . . .                                                                                                                                                                                                                                      | 170 |
| Table 5.2:  | Comparison of the LM-SR1 and BFGS quasi-Newton algorithms for<br>CR-CC(2,3),D/TZVP geometry optimizations. <sup>a</sup> . . . . .                                                                                                                                                                                                                            | 174 |
| Table 5.3:  | Details of the wall times (hours) characterizing the CCSD(T)/SBKJC(1f)<br>parallel numerical optimization of the Au <sub>8</sub> S <sub>1</sub> isomer employing the<br>D <sub>2h</sub> symmetry (4 degrees of freedom), performed using nine 8-core<br>nodes <sup>a</sup> . . . . .                                                                         | 184 |
| Table 5.4:  | Bond lengths (in Å) of the S <sub>1</sub> , S <sub>3</sub> , S <sub>4</sub> , and S <sub>6</sub> isomers of Au <sub>8</sub> ob-<br>tained from geometry optimizations using MP2, CCSD(T), and some<br>representative DFT approaches (see Fig. 5.1 for the meaning of the<br>geometrical parameters). . . . .                                                 | 186 |
| Table 5.5:  | Relative energies (in kcal/mol) of the S <sub>1</sub> , S <sub>3</sub> , S <sub>4</sub> , and S <sub>6</sub> isomers<br>of Au <sub>8</sub> , with respect to the S <sub>1</sub> isomer, obtained in the MP2 and<br>CCSD(T) calculations. . . . .                                                                                                             | 190 |
| Table 5.6:  | Comparison of the energies (in millihartree) of RHF- and UHF-based<br>CCSD and various triples-corrected CC approximations with the cor-<br>responding full CI data <sup>a</sup> for the equilibrium and four displaced ge-<br>ometries ( <i>R<sub>e</sub></i> = 1.7328 bohr) of the HF molecule, as described by the<br>spherical Dunning DZ basis. . . . . | 198 |
| Table 5.7:  | Comparison of the energies (in millihartree) of RHF- and UHF-based<br>CCSD and various triples-corrected CC approximations with their<br>parent CCSDT approach data for the equilibrium and four displaced<br>geometries ( <i>R<sub>e</sub></i> = 1.7328 bohr) of the HF molecule, as described by<br>the spherical Dunning DZ basis. . . . .                | 199 |

|             |                                                                                                                                                                                                                                                                                                                  |     |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 5.8:  | Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with the corresponding full CI data for the equilibrium and displaced geometries of the F <sub>2</sub> molecule, as described by the spherical cc-pVDZ basis set.                        | 202 |
| Table 5.9:  | Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the F <sub>2</sub> molecule, as described by the spherical cc-pVDZ basis set.                           | 203 |
| Table 5.10: | Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the H <sub>2</sub> O <sub>2</sub> molecule, <sup>a</sup> as described by the spherical cc-pVDZ basis.   | 207 |
| Table 5.11: | Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the CH <sub>3</sub> CH <sub>3</sub> molecule, <sup>a</sup> as described by the spherical cc-pVDZ basis. | 210 |

## LIST OF FIGURES

|             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |    |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 3.1: | Orbital levels of the $\text{CH}^+$ and $\text{OH}^-$ ions and a schematic representation of the electron attachment and ionization processes that lead to the formation of the CH and OH radicals from the $\text{CH}^+$ and $\text{OH}^-$ reference closed-shell systems. Valence shells of $\text{CH}^+$ and $\text{OH}^-$ that play a dominant role in the relevant electron attachment and ionization processes and that are used in the active-space EA- and IP-EOMCC calculations are emphasized with the help of dotted frames. . . . .                                                                                                                                                                                                                                                                                                                                                                               | 18 |
| Figure 3.2: | Vertical IEs of $\text{Au}_3^-$ (red bars) superimposed on the photoelectron spectrum from Fig. 1(c) in Ref. [222]: (a) IP-EOMCCSD( $2h-1p$ )/aug-cc-pVDZ-PP calculations; (b) IP-EOMCCSD( $2h-1p$ )/aug-cc-pVTZ-PP+CV calculations; (c) IP-EOMCCSD( $3h-2p$ )/aug-cc-pVDZ-PP calculations; (d) extrapolated IEs determined using Eq. (3.21). For the actual symmetries and energies of the calculated states of $\text{Au}_3$ , see Table 3.8. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 68 |
| Figure 4.1: | The $C_{2v}$ -symmetric structure of azulene and its bond lengths. The large black spheres represent carbon atoms, whereas the small gray spheres represent hydrogens. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 84 |
| Figure 4.2: | The photoelectron spectrum of azulene using a probe pulse centered at 400 nm, recorded for four pump excitation energies (the corresponding vibrational energies in $S_2$ given in parentheses), namely, 201 nm (2.61 eV; Ref. [142]), 268 nm (1.07 eV; Ref. [243]), 283 nm (0.82 eV; Ref. [243]), and 335 nm (0.14 eV; Ref. [243]), and a pump-probe delay time of 500 fs (for other pump-probe delay times, which yield similar spectra, see Ref. [243]). The unstructured direct ionization (65 % of photoionization events) has been subtracted. Each spectral profile represents a fingerprint of the Rydberg states from which azulene is photoionized after absorbing the second probe photon. The electronic energies of these Rydberg states are marked by $R_A$ , $R_B$ , etc. Two horizontal axes near the top show the vibrational energies in Rydberg states for the pump wavelengths of 201 and 335 nm. . . . . | 88 |

|             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |     |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4.3: | The proposed schematics of the $1 + 2'$ photoionization experiment [243]. The Rydberg states, from which azulene is photoionized, are populated by the electronic relaxation from the postulated doubly excited state located below the ionization threshold $D_0$ , marked by **. Excitation of the doubly excited state from $S_2$ , relaxation into the Rydberg states, and photoionization take place within the probe pulse duration (100 fs). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 89  |
| Figure 4.4: | Schematic representation of the reaction pathway I for the methanol oxidation to formic acid on the $\text{Au}_8^-$ cluster proposed in Ref. [300].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 107 |
| Figure 4.5: | Energy diagram for reaction pathway I characterizing the oxidation of methanol to formic acid on the $\text{Au}_8^-$ particle resulting from CR-CC(2,3),D (green bold font), M06 (red italic font), B3LYP (black roman font), and $\omega$ -B97X-D (blue roman font) calculations using the BS2 and, in parentheses, BS1 basis sets. The CR-CC(2,3),D/BS1 energies and all of the DFT energies represent truly computed data. The CR-CC(2,3),D/BS2 energies were obtained by extrapolation based on Eq. (4.28), in which the difference between the CCSD energies obtained with the BS2 and BS1 basis sets is added to the CR-CC(2,3),D/BS1 energy. The energies shown for the A1, B1, C1, and D1 intermediates are relative to the $\text{CH}_3\text{O}^- + \text{O}_2 + \text{Au}_8^-$ reactants. The energies shown for the products are relative to the D1 intermediate. The energies shown for the transition states TS_A1, TS_B1, and TS_C1 are the corresponding (TS_A1 – A1), (TS_B1 – B1), and (TS_C1 – C1) barrier heights. All energies are in kcal/mol. Also shown are the molecular structures defining the stationary points along pathway I resulting from the M06 geometry optimizations carried out in Ref. [300]. . . . . | 117 |
| Figure 4.6: | Electron distributions among the nearly degenerate HOMO and LUMO making up the four lowest-energy electronic states of $\text{C}_4\text{O}_4$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 139 |
| Figure 4.7: | Orbital classification used in the active-space CC approaches, such as CCSDt. Full circles represent core and active electrons of the reference determinant $ \Phi\rangle$ , which is the formal reference defining the Fermi vacuum for the active-space CC calculations. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 144 |
| Figure 5.1: | The $S_1$ , $S_3$ , $S_4$ , and $S_6$ isomers of $\text{Au}_8$ examined in this study, along with the selected geometrical parameters included in Table 5.4. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 181 |

|             |                                                                                                                                                                                                                                                                                                                                                             |     |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 5.2: | Energy differences of the various RHF- and UHF-based coupled-cluster approaches with some form of triples with respect to the full CI results for (a) the HF molecule and (b) the F <sub>2</sub> diatomic species at various bond lengths. . . . .                                                                                                          | 204 |
| Figure 5.3: | Energy differences of the various RHF- and UHF-based approximate triples coupled-cluster methods with respect to their parent CCSDT results for (a) the HF molecule, (b) the F <sub>2</sub> diatomic species, (c) the H <sub>2</sub> O <sub>2</sub> species, and (d) the C <sub>2</sub> H <sub>6</sub> polyatomic molecule at various bond lengths. . . . . | 212 |

# Chapter 1

## Introduction

Coupled-cluster (CC) theory [1–6], which is known to offer the best compromise between computational cost and accuracy and which results in approximations that lead to a rapid convergence to the exact, full configuration interaction (CI) limit, has become the *de facto* standard for high-accuracy molecular electronic structure calculations for ground and excited states and properties other than energy (see Refs. [7–11] for reviews). However, in order for the ground-state CC methods and their various excited-state extensions to be successful, by providing chemical or near-spectroscopic accuracies, one needs to find computationally efficient and robust treatments of higher-than-double excitations, such as triply or triply and quadruply excited clusters. Indeed, the basic CC method with singles and doubles (CCSD) [12–15], which has relatively inexpensive CPU steps that scale as  $n_o^2 n_u^4$  or  $\mathcal{N}^6$  ( $n_o$  and  $n_u$  are, respectively, the numbers of occupied and unoccupied orbitals used in the correlated calculations and  $\mathcal{N}$  is a measure of the system size), although generally more accurate than its CI (i.e., CISD) counterpart, especially in larger systems, where the lack of size-extensivity of CISD becomes a major problem, is insufficient in obtaining accurate ground-state properties, especially energetics (reaction energies, activation barriers, chemical reaction profiles, etc.). At the same time, the extension of CCSD to excited states via the equation-of-motion CCSD (EOMCCSD) approach [16–18] and its symmetry-adapted-cluster (SAC) CI [19–22] and

linear-response CC [23–28] counterparts, although useful in a qualitative characterization of excited states dominated by one-electron transitions, is generally not accurate enough to obtain a quantitative description of such states, especially when larger polyatomic species are examined (cf., e.g., Refs. [29–32]; for a thorough evaluation of a number of EOMCC methods, including EOMCCSD, illustrating the same, see Refs. [33–39]). EOMCCSD and its SAC-CI and linear-response counterparts also fail when the electronically excited states of interest are characterized by significant two-electron or higher many-electron transitions [39–54]. When the connected triple excitations are explicitly included in the CC and EOMCC considerations via the CCSDT approach [55, 56] and its excited-state EOMCCSDT extension [42, 43, 57], the description of the ground-state energetics and properties and excited states significantly improves, yielding often the virtually exact results (see, e.g., Refs. [40–43, 54–61]). However, it is also accompanied by a steep increase in the iterative CPU time scaling and excitation amplitude storage requirements characterizing the CCSDT/EOMCCSDT approximation, from  $n_o^2 n_u^4$  ( $\mathcal{N}^6$ ) and  $n_o^2 n_u^2$  ( $\mathcal{N}^4$ ) in the CCSD/EOMCCSD case to  $n_o^3 n_u^5$  ( $\mathcal{N}^8$ ) and  $n_o^3 n_u^3$  ( $\mathcal{N}^6$ ), respectively, in the case of CCSDT/EOMCCSDT, limiting its applicability to systems with up to a dozen or so correlated electrons and smaller basis sets. Thus, if one is to make use of the CC/EOMCC methodologies in accurate molecular electronic structure calculations, the CC/EOMCC schemes that can account for the effects of triples in an approximate, cost effective, and yet reliable manner need to be employed.

One way the effects of triple and other higher-than-double excitations can be included in the CC/EOMCC approaches without having to deal with the prohibitive computational costs of the full CCSDT/EOMCCSDT methods is through the use of active orbitals, as done



in the ground-state CCSDt [62–67] and excited state EOMCCSDt [42–44] approaches (cf. Refs. [54, 68] for reviews). While this allows for full CCSDT/EOMCCSDT-quality results at the cost of CCSD/EOMCCSD times a prefactor proportional to the numbers of active occupied and active unoccupied orbitals used to select the triples, the approach is no longer, strictly speaking, a pure computational black box as one has to select the active orbitals involved in identifying the dominant triples. The same applies to higher-order methods in this category, such as CCSDtq [54, 63–66, 69].

One can also contemplate approaches for identifying the most important triples (and higher) contributions through the many-body perturbation theory (MBPT) analysis, as is commonly done when developing approximate quasi-perturbative CC/EOMCC approximations, best represented by the oldest methods in this category, such as CCSD[T] [70, 71], CCSD(T) [72], and CCSDT-1 [71, 73]. Methods of this type are computational black boxes and reduce computer costs of their parent full approximations by orders of magnitude, while providing an accurate description of molecules near the equilibrium geometries. This is best symbolized by the success of the CCSD(T) approach, which replaces the iterative  $n_o^3 n_u^5$  CPU steps of CCSDT by the iterative  $n_o^2 n_u^4$  CPU steps of CCSD and the noniterative  $n_o^3 n_u^4$  steps associated with the determination of the triples correction to the CCSD energy, while offering the CCSDT level of accuracy in the ground-state calculations as long as one does not significantly stretch or break chemical bonds. Problems emerge, however, when one wants to apply methods of the CCSD(T) type to bond breaking or biradicals, or consider excited states, especially those having the more substantial contributions from two-electron transitions (cf., e.g., Refs. [39–44, 54, 66–68]). In response to these challenges, several alternatives or extensions

of CCSD(T) that improve the CCSD(T) results in bond breaking and other multireference situations and that improve EOMCCSD results for excited states without significantly increasing costs and ease of use of CCSD(T) calculations have been formulated over the years. Some examples of these alternatives to CCSD(T) or robust extensions of CCSD(T)-like approaches to excited states include the noniterative triples or triples and quadruples methods for the ground-state computations defined the CCSD(2) or CC(2)PT(2) and CCSD(2)<sub>T</sub> approaches [74–77] (see Refs. [78–80] for related ideas), and their excited-state analogs in the form of the EOMCC(2)PT(2) method [76] and its size-intensive EOMCCSD(2)<sub>T</sub> modification [45], the linear-response CCSDR(3) scheme [81, 82], and its iterative, still rather inexpensive CC3 parent [81–84], the EOMCCSD(T) [85], EOMCCSD( $\tilde{T}$ ) [86], and EOMCCSD(T') [86] hierarchy obtained from the perturbative analysis of the EOMCCSDT equations, and a wide variety of noniterative corrections to CCSD or EOMCCSD energies resulting from the method of moments of CC (MMCC) equations [40, 41, 46–52, 87–95], such as CR-CC(2,3) [51, 93–95], CR-EOMCC(2,3) [51, 52], and the recently implemented  $\delta$ -CR-EOMCC(2,3) scheme [31]. The MMCC approaches, such as CR-CC(2,3), CR-EOMCC(2,3), and  $\delta$ -CR-EOMCC(2,3) are particularly promising, since they retain the black-box nature of the popular ground-state CCSD(T) approximation and reduce the large costs of the full CCSDT/EOMCCSDT calculations to the more manageable iterative  $n_o^2 n_u^4$  ( $\mathcal{N}^6$ ) and non-iterative  $n_o^3 n_u^4$  ( $\mathcal{N}^7$ ) computational steps, while providing a highly accurate description of chemical reaction pathways involving single bond breaking, biradicals, and excited states dominated by one- and two-electron transitions [39, 44, 46–51, 51–53, 93, 94, 96–99]. The CR-CC(2,3) and related approaches are capable of producing highly accurate results for a vast

array of molecular systems with up to about 100 correlated electrons, when the canonical formulation is employed, or hundreds or even thousands of electrons, when the local correlation ideas are implemented [99–103].

While the CR-CC and CR-EOMCC approaches have demonstrated considerable success (cf., e.g., Refs. [29–31, 40, 41, 48, 49, 51, 51–53, 88, 89, 93, 94, 96–99, 104–149]), they are not free from all the problems. There are, for example, cases of biradical reaction mechanisms, where they do not work as well as desired, resulting in the underestimation of electronic energies of singlet biradicals, due to significant coupling of singles, doubles, and connected triples in such situations, which the CR-CC(2,3) and other noniterative corrections to CCSD neglect. This issue can be addressed by combining the aforementioned active-space CC/EOMCC (e.g., CCSDt) and CR-CC/CR-EOMCC (e.g., CR-CC(2,3)) ideas in the form of the so-called  $CC(P; Q)$  methodology [68, 150, 151]. The resulting  $CC(t,3)$  approach combining CCSDt iterations with noniterative corrections due to certain categories of triples missing in CCSDt provides spectacular results in describing bond breaking, biradical reaction mechanisms, and singlet-triplet gaps in biradicals [68, 150, 151]. There is, however, another problem, which none of the above methods addresses in a satisfactory manner, namely the issue of proper spin adaptation, which is difficult to achieve in conventional CC theory when applied to open-shell states. For example, the usual open-shell implementations of the CR-CC(2,3) and CR-EOMCC(2,3) and other CC/EOMCC schemes, which utilize a spin-integrated spin-orbital formalism in derivations and computer implementations and which make use of either the unrestricted or restricted open-shell Hartree-Fock (UHF or ROHF, respectively) reference determinants, do not properly account for the spin symmetry when electronic states of

interest are not singlets, and, as a result, may suffer from the spin-contamination problem (cf., e.g., Ref. [152]) or other issues, such as the non-analytic behavior and substantial loss of accuracy in the region of Hartree-Fock instabilities [153].

Issues related to spin-adaption of CC/EOMCC calculations for open-shell systems, such as radicals and biradicals, can be addressed by turning to the alternative hierarchy of EOMCC approaches, termed the electron-attached (EA) [154–162] and ionized (IP) [157–169] EOMCC theories as well as their analogous doubly attached (DEA) and doubly ionized (DIP) approaches [170–178]. In particular, the EA- and IP-EOMCC methods enable one to perform orthogonally spin-adapted calculations for the ground and excited states of the  $(N\pm 1)$ -electron open-shell systems around  $N$ -electron closed shells by applying the linear electron-attaching or ionizing operator,  $R_\mu^{(N\pm 1)}$ , to the correlated ground state of the reference  $N$ -electron closed-shell core obtained with the single-reference CC approach. Because these methods use a closed-shell reference determinant and because one obtains target  $(N\pm 1)$ -electron states by diagonalizing the similarity transformed Hamiltonian obtained in the closed-shell CC calculations for the underlying  $N$ -electron system, which commutes with  $S^2$  and  $S_z$ , they provide a convenient formalism for performing orthogonally spin- and symmetry-adapted calculations for radicals and ions of closed-shell species, eliminating the issues associated with spin-contamination and making the interpretation of the calculated electronic states much more transparent than in the traditional ROHF or UHF based CC/EOMCC calculations. In analogy to our earlier discussed CCSD/EOMCCSD or CCSDT/EOMCCSDT, the basic low-order EA- and IP-EOMCC approximations including up to 2-particle-1-hole ( $2p-1h$ ) and 2-hole-1-particle ( $2h-1p$ ) excitations, here referred to as

EA-EOMCCSD( $2p-1h$ ) and IP-EOMCCSD( $2h-1p$ ), although useful in calculations of the electron affinities and ionization potentials of closed-shell molecules, respectively, are insufficient to provide an accurate description for the ground and, especially, excited states of radicals and similar open-shell systems [154, 157–160, 179]. The inclusion of higher-order components of  $R_\mu^{(N\pm1)}$ , such as 3-particle-2-hole ( $3p-2h$ ) and 3-hole-2-particle ( $3h-2p$ ), excitations, greatly improves the results, but also significantly increases the computational costs, limiting the use of the resulting EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ) approaches and their EA- and IP-EOMCCSDT analogs [156, 167, 168] to small molecular systems. To remedy this problem, the aforementioned active-space CC and EOMCC methodologies [42–44, 54, 62–69] methodologies have been extended to the EA-EOMCC and IP-EOMCC approaches [157–161]. The active-space EA- and IP-EOMCC schemes use only a small subset of all higher than  $2p-1h$  and higher than  $2h-1p$  excitations, respectively, which are chosen through a suitably defined set of active orbitals. In analogy to particle-conserving active-space CC/EOMCC methods, such as CCSDt or EOMCCSDt, this significantly reduces the computational costs associated with inclusion of all higher-order excitations in the EA- and IP-EOMCC approaches, but since the most important higher-order contributions are still accounted for, the active-space EA- and IP-EOMCC methods are capable of producing highly accurate results of the same quality as their parent approximations. The EA-EOMCC and IP-EOMCC approaches with an active-space treatment of  $3p-2h$  and  $3h-2p$  excitations [157–159], designated here and elsewhere in this document as the EA-EOMCCSDt $\{N_u\}$  and IP-EOMCCSDt $\{N_o\}$  methods, where  $N_o$  and  $N_u$  are the number of active occupied and active unoccupied orbitals, respectively, have demonstrated their reliabil-

ity in reproducing the results for the ground and excited states of radicals obtained with their more expensive parent approximations, EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ), respectively, and other high-level *ab initio* approaches at a small fraction of the computer cost [157–160,180]. Similar remarks apply to the recently developed DEA- and DIP-EOMCC methods with an active-space treatment of  $4p-2h$  and  $4h-2p$  excitations, which are particularly useful in examining the electronic spectra of biradical species [177,178].

In this dissertation, we have used the EA- and IP-EOMCC approaches with full and active-space treatments of  $3p-2h$  and  $3h-2p$  excitations to investigate several challenging situations involving small and yet complicated open-shell molecular species, such as the adiabatic excitation spectra of the CNC, C<sub>2</sub>N, N<sub>3</sub>, and NCO molecules [180] and the photoelectron spectra of Au<sub>*n*</sub><sup>−</sup> particles, demonstrating the robustness and utility of these approaches in such challenging cases. In addition to the electronic structure of these systems, we have used parallel numerical derivatives for the EA- and IP-EOMCC methods, developed in this PhD research, to examine the ability of the EA/IP-EOMCC methods with up to  $3p-2h/3h-2p$  excitations to provide accurate nuclear configuration information, including excited states. For several other closed- and open-shell species, we have used the CR-CC(2,3) and  $\delta$ -CR-EOMCC(2,3) approaches and parallel numerical derivatives based on the CR-CC(2,3) and CR-EOMCC(2,3) levels, developed in this thesis work too, to solve important chemical and spectroscopic problems, providing definitive confirmation of the existence of the highly-correlated doubly excited state of azulene below the ionization threshold, which drives the  $1 + 2'$  multiphoton ionization experiments that lead to a clear Rydberg fingerprint [142], examining the details of the mechanism and energetics of the aerobic oxidation

of methanol on  $\text{Au}_8^-$  nanoparticle [181], and obtaining a highly accurate description of the low-lying states of the deceptively simple 1,2,3,4-cyclobutanetetraone [182] confirming precise photodetachment experiments performed recently. As already alluded to above, as part of these investigations we have developed and implemented parallel numerical derivative routines which can be used with any molecular electronic structure method (including and CC/EOMCC approach) in conjunction with optimization algorithm routines to help speed up the evaluation of geometries and harmonic vibrational frequencies. This has allowed us to provide definitive information about the geometries and relative energies of the low-energy isomers of the controversial  $\text{Au}_8$  particle at the high CCSD(T) level [183]. Finally, returning to the issue of UHF *vs* RHF (restricted Hartree-Fock) reference determinant in CC calculations, we have examined the effects of using spin-symmetry broken UHF rather than spin-adapted RHF reference wave function on the CR-CC(2,3) results in describing bond-breaking on singlet potential energy surfaces, showing that the spin-adapted closed-shell CR-CC(2,3) codes based on RHF references provide the best overall and most accurate results. In summary, the focus of this dissertation has been the application of novel high-level CC/EOMCC approaches developed in our group to challenging and experimentally relevant chemical situations, considered to be highly complex for conventional quantum chemistry, and the development of new codes and algorithms for high-level geometry optimizations and frequency calculations and use of unrestricted references in CR-CC computations.

# Chapter 2

## Project objectives

The main objectives of this work are:

- A. Application of the full and active-space EA- and IP-EOMCC methods to the energies and geometries of the ground and excited states of CNC, C<sub>2</sub>N, NCO, and N<sub>3</sub>. Interpretation of the photoelectron spectrum of Au<sub>3</sub><sup>-</sup> using the IP-EOMCC methodologies.
- B. Application of the CR-CC(2,3) and  $\delta$ -CR-EOMCC(2,3) approaches to highly-correlated systems, including the doubly excited state of azulene below the ionization threshold mediating the 1+2' multiphoton ionization, the aerobic oxidation of methanol to formic acid on Au<sub>8</sub><sup>-</sup>, and the low-lying electronic states of 1,2,3,4-cyclobutanetetraone.
- C. Development and testing of parallel numerical gradients and second derivatives for geometry optimizations and vibrational harmonic frequencies using various CC/EOMCC approximations, including CCSD(T), CR-CC(2,3), and CR-EOMCC(2,3), and the EA/IP-EOMCC methods with up to 3*p*-2*h* and 3*h*-2*p* excitations treated fully or via active orbitals. Examining alternative algorithms for accelerating the convergence of the geometry optimization methods using numerically determined energy derivatives.
- D. Implementation and benchmarking of unrestricted CR-CC(2,3).



# Chapter 3

## Applications of full and active-space electron attached and ionized equation-of-motion coupled-cluster methods

In this chapter, we discuss the application of the EA- and IP-EOMCC approaches with up to  $3p-2h$  and  $3h-2p$  excitations, treated fully or with active orbitals, to challenging open-shell problems. Section 3.1 provides the theoretical background of the relevant EA- and IP-EOMCC methodologies. Section 3.2 discusses their application to a series of challenging triatomic radicals, CNC, C<sub>2</sub>N, N<sub>3</sub>, and NCO, including geometry optimizations and adiabatic excitation energies, based on the results reported in Ref. [180]. Finally, in Section 3.3, we demonstrate the utility of the IP-EOMCC approaches in providing an accurate interpretation of not only peak positions, but also peak widths and intensities, for the previously unexplained photoelectron spectrum of Au<sub>3</sub><sup>-</sup>, relying on our calculations reported in Ref. [184].

### 3.1 Theory

In the EA/IP-EOMCC theories, one obtains the ground ( $\mu = 0$ ) and excited ( $\mu > 0$ ) states  $|\Psi_\mu^{(N\pm 1)}\rangle$  of an  $(N \pm 1)$ -electron system by applying a linear electron-attaching or ionizing operator  $R_\mu^{(N\pm 1)}$  to the CC ground state

$$|\Psi_0^{(N)}\rangle = e^T |\Phi\rangle \quad (3.1)$$

of the related closed-shell  $N$ -electron system, so that

$$|\Psi_\mu^{(N\pm 1)}\rangle = R_\mu^{(N\pm 1)} |\Psi_0^{(N)}\rangle \equiv R_\mu^{(N\pm 1)} e^T |\Phi\rangle. \quad (3.2)$$

Here,

$$T = \sum_{n=1}^N T_n, \quad (3.3)$$

where

$$T_n = \sum_{i_1 < \dots < i_n, a_1 < \dots < a_n} t_{a_1 \dots a_n}^{i_1 \dots i_n} a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1}, \quad (3.4)$$

is the cluster operator of the ground-state CC theory obtained by solving the CC equations for the  $N$ -electron reference system, with  $t_{a_1 \dots a_n}^{i_1 \dots i_n}$ , representing the corresponding cluster amplitudes,  $|\Phi\rangle$  is the  $N$ -electron reference determinant (typically, the RHF configuration), and the  $R_\mu^{(N\pm 1)}$  operators entering Eq. (3.2) are defined as

$$R_\mu^{(N+1)} = \sum_{n=0}^N R_{\mu, (n+1)p-nh}, \quad (3.5)$$

where

$$R_{\mu,(n+1)p-nh} = \sum_{i_1 < \dots < i_n, a < a_1 < \dots < a_n} r_{aa_1 \dots a_n}^{i_1 \dots i_n} a^a a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1} \quad (3.6)$$

in the EA case, and

$$R_{\mu}^{(N-1)} = \sum_{n=0}^N R_{\mu,(n+1)h-np}, \quad (3.7)$$

where

$$R_{\mu,(n+1)h-np} = \sum_{i_1 < \dots < i_n < i, a_1 < \dots < a_n} r_{a_1 \dots a_n}^{ii_1 \dots i_n} a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1} a_i \quad (3.8)$$

in the IP case. Here and elsewhere in this dissertation,  $i, j, \dots$  ( $a, b, \dots$ ) are the spin-orbitals occupied (unoccupied) in the reference determinant  $|\Phi\rangle$  and  $a^p$  ( $a_p$ ) are the usual creation (annihilation) operators associated with the spin-orbitals  $\{|p\rangle\}$ .

Once the ground-state CC equations are solved for the cluster amplitudes  $t_{a_1 \dots a_n}^{i_1 \dots i_n}$  and the ground-state CC energy  $E_0^{(N)}$  of the  $N$ -electron reference system is determined, one obtains the  $1p$  amplitudes  $r_a$ , the  $2p-1h$  amplitudes  $r_{ab}^j$ , the  $3p-2h$  amplitudes  $r_{abc}^{jk}$ , etc. of the EA-EOMCC theory entering Eq. (3.5) or their IP-EOMCC  $1h$ ,  $2h-1p$ ,  $3h-2p$ , etc. analogs entering Eq. (3.7), and the corresponding electron-attachment or ionization energies

$$\omega_{\mu}^{(N\pm 1)} = E_{\mu}^{(N\pm 1)} - E_0^{(N)}, \quad (3.9)$$

where  $E_{\mu}^{(N+1)}$  and  $E_{\mu}^{(N-1)}$  are the energies of ground and excited states of the target  $(N+1)$ - and  $(N-1)$ -electron systems, respectively, by diagonalizing the similarity-transformed Hamiltonian, written here in the normal-ordered form relative to the Fermi vacuum  $|\Phi\rangle$ ,

$$\bar{H}_{N,\text{open}} = (H_N e^T)_{C,\text{open}} = e^{-T} H_N e^T - (H_N e^T)_{C,\text{closed}} \quad (3.10)$$

where the subscripts “open”, “closed”, and  $C$  refer to the open (i.e., having external lines), closed (i.e., having no external lines), and connected parts of a given operator expression, in the appropriate subspace of the  $(N \pm 1)$ -electron Hilbert-space spanned by the determinants corresponding to the many-body components included in  $R_\mu^{(N \pm 1)}$ . Thus, in the EA-EOMCCSD( $3p-2h$ ) method, which interests us in this work and in which

$$T = T_1 + T_2 \quad (3.11)$$

is obtained in the ground-state CCSD calculations for the  $N$ -electron reference system and

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p-1h} + R_{\mu,3p-2h}, \quad (3.12)$$

we diagonalize the similarity-transformed Hamiltonian of CCSD,

$$\bar{H}_{N,\text{open}}^{(\text{CCSD})} = (H_N e^{T_1 + T_2})_{C,\text{open}}, \quad (3.13)$$

in the subspace of the  $(N + 1)$ -electron Hilbert-space spanned by the  $|\Phi^a\rangle = a^a|\Phi\rangle$ ,  $|\Phi_j^{ab}\rangle = a^a a^b a_j|\Phi\rangle$ , and  $|\Phi_{jk}^{abc}\rangle = a^a a^b a^c a_k a_j|\Phi\rangle$  determinants. In the IP-EOMCCSD( $3h-2p$ ) approach, which interests us in this study as well, where  $T$  is again obtained in the CCSD calculations and

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h-1p} + R_{\mu,3h-2p}, \quad (3.14)$$

we diagonalize  $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$  in the subspace of the  $(N - 1)$ -electron Hilbert-space spanned by the  $|\Phi_i\rangle = a_i|\Phi\rangle$ ,  $|\Phi_{ij}^b\rangle = a^b a_j a_i|\Phi\rangle$ , and  $|\Phi_{ijk}^{bc}\rangle = a^b a^c a_k a_j a_i|\Phi\rangle$  determinants. In the basic

EA-EOMCCSD( $2p-1h$ ) and IP-EOMCCSD( $2h-1p$ ) approaches used in our calculations too, we neglect the  $3p-2h$  component  $R_{\mu,3p-2h}$  in Eq. (3.12) and the  $3h-2p$  component  $R_{\mu,3h-2p}$  in Eq. (3.14), and diagonalize  $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$  in the smaller subspaces spanned by the  $|\Phi^a\rangle$  and  $|\Phi_j^{ab}\rangle$  determinants in the EA case and  $|\Phi_i\rangle$  and  $|\Phi_{ij}^b\rangle$  determinants in the IP case.

It has been well established (largely by our group [157–160]) that inclusion of higher-order terms in the  $R_{\mu}^{(N\pm 1)}$  operators, especially the  $R_{\mu,3p-2h}$  component in  $R_{\mu}^{(N+1)}$  and the  $R_{\mu,3h-2p}$  component in  $R_{\mu}^{(N-1)}$ , is necessary to obtain a quantitative description of the electronic excitations in radical species. The EA-EOMCCSD( $2p-1h$ ) and IP-EOMCCSD( $2h-1p$ ) methods, abbreviated sometimes as EA-EOMCCSD and IP-EOMCCSD, work well for the lowest electron attachment and ionization energies, but they usually fail otherwise, especially when electronic spectra of radicals need to be examined. The results of the basic EA-EOMCCSD( $2p-1h$ ) and IP-EOMCCSD( $2h-1p$ ) calculations become particularly poor when the  $2p-1h$  and  $2h-1p$  contributions relative to the  $N$ -electron reference  $|\Phi\rangle$  or the two-electron ( $2p-2h$ ) transitions relative to the ground state of the radical of interest become significant. Unfortunately, the full inclusion of the  $R_{\mu,3p-2h}$  and  $R_{\mu,3h-2p}$  terms in the EA- and IP-EOMCC calculations comes at a high price, increasing the  $\mathcal{N}^5$ -like  $n_o n_u^4$  and  $n_o^2 n_u^3$  operations defining the iterative diagonalization steps of EA-EOMCCSD( $2p-1h$ ) and IP-EOMCCSD( $2h-1p$ ) to the  $\mathcal{N}^7$ -like  $n_o^2 n_u^5$  and  $n_o^3 n_u^4$  steps. One way to retain the accuracy of the higher-order EA- and IP-EOMCC schemes with the  $3p-2h$  and  $3h-2p$  excitations, while avoiding this steep computer cost increase, is to use the active-space variants of the EA/IP-EOMCC methods developed in the Piecuch research group, described in Refs. [54, 157–160] and summarized below.

In the active-space EA/IP-EOMCC approaches, one defines a suitable set of active orbitals, which are used to *a priori* select a small number of the most important higher-than- $2p$ - $1h$  contributions to  $R_\mu^{(N+1)}$  and higher-than- $2h$ - $1p$  contributions to  $R_\mu^{(N-1)}$ . This is based on the observation that in many cases one can envision the formation of a radical by either attaching an electron to one of the lowest-energy unoccupied orbitals or by removing an electron from one of the highest-energy occupied orbitals of the related closed-shell system. (see Fig. 3.1 for an illustration of the formation of the CH and OH radicals from the underlying closed-shell systems,  $\text{CH}^+$  and  $\text{OH}^-$ , respectively. It is, therefore, reasonable, in a manner similar to that of multireference approaches, to use these orbitals as active orbitals in the EA- and IP-EOMCC calculations. Thus, we divide the available spin-orbitals of a closed-shell  $N$ -electron system into four disjoint groups of core or inactive occupied spin-orbitals, labeled by lower-case bold letters **i**, **j**,  $\dots$ , active spin-orbitals occupied in the reference determinant  $|\Phi\rangle$ , labeled by upper-case bold letters **I**, **J**,  $\dots$ , active spin-orbitals unoccupied in  $|\Phi\rangle$ , labeled by upper-case bold letters **A**, **B**,  $\dots$ , and virtual or inactive unoccupied spin-orbitals, labeled by lower-case bold letters **a**, **b**,  $\dots$ . We continue to designate the occupied and unoccupied spin-orbitals in  $|\Phi\rangle$  by the italic characters  $i$ ,  $j$ ,  $\dots$  and  $a$ ,  $b$ ,  $\dots$ , respectively, if the active/inactive character of the spin-orbitals is not specified. Once the above orbital classification scheme is established, we use it to define the electron attaching and electron removing operators  $R_\mu^{(N\pm 1)}$  of the active-space EA- and IP-EOMCC and SAC-CI methods, following the general recipe introduced in Refs. [157–162] (see also, Ref. [54]). In particular, the active-space EA-EOMCCSD( $3p$ - $2h$ ) $\{N_u\}$  approach using  $N_u$  active unoccupied orbitals is obtained by replacing the  $3p$ - $2h$  component  $R_{\mu,3p-2h}$  of the electron attaching operator

$R_\mu^{(N+1)}$ , Eq. (3.12), by [54, 157–160]

$$r_{\mu,3p-2h} = \sum_{j>k, \mathbf{A}<b<c} r_{\mathbf{A}bc}^{jk} a^{\mathbf{A}} a^b a^c a_k a_j. \quad (3.15)$$

The relatively small set of the unknown amplitudes  $r_{\mathbf{A}bc}^{jk}$  defining  $r_{\mu,3p-2h}$ , Eq. (3.15), in which at least one of the three unoccupied spin-orbital indices is active, and the remaining  $1p$  and  $2p-1h$  amplitudes,  $r_a$  and  $r_{ab}^j$ , respectively, that enter the  $(N+1)$ -electron wave functions of the active-space EA-EOMCCSD( $3p-2h$ ){ $N_u$ } approach are obtained by diagonalizing the similarity-transformed Hamiltonian of CCSD, Eq. (3.13), in the subspace of the  $(N+1)$ -electron Hilbert space spanned by the  $|\Phi^a\rangle$ ,  $|\Phi_{j'}^{ab}\rangle$ , and  $|\Phi_{jk}^{\mathbf{A}bc}\rangle$  determinants.

This reduces the iterative  $n_o^2 n_u^5$  steps of the parent EA-EOMCCSD( $3p-2h$ ) approach to the  $N_u n_o^2 n_u^4$  steps, which typically are a small prefactor  $N_u$  ( $\ll n_u$ ) times the cost of the CCSD calculations, enabling one to perform the EA-EOMCCSD( $3p-2h$ )-level calculations for much larger systems and basis sets than the full EA-EOMCCSD( $3p-2h$ ) approach using all  $3p-2h$   $r_{abc}^{jk}$  amplitudes would normally allow. Similarly, the active-space IP-EOMCCSD( $3h-2p$ ){ $N_o$ } approach using  $N_o$  active occupied orbitals is obtained by replacing the  $3h-2p$  component  $R_{\mu,3h-2p}$  of the ionizing operator  $R_\mu^{(N-1)}$ , Eq. (3.14), by [54, 157–160]

$$r_{\mu,3h-2p} = \sum_{\mathbf{I}>j>k, b<c} r_{bc}^{\mathbf{I}jk} a^{\mathbf{I}} a^b a^c a_k a_j a_{\mathbf{I}}, \quad (3.16)$$

where the relatively small set of the unknown amplitudes  $r_{bc}^{\mathbf{I}jk}$  defining  $r_{\mu,3h-2p}$ , Eq. (3.16), in which at least one of the three occupied spin-orbital indices is active, and the remaining  $1h$  and  $2h-1p$  amplitudes,  $r^i$  and  $r_b^{ij}$ , respectively, that define the  $(N-1)$ -electron wave functions

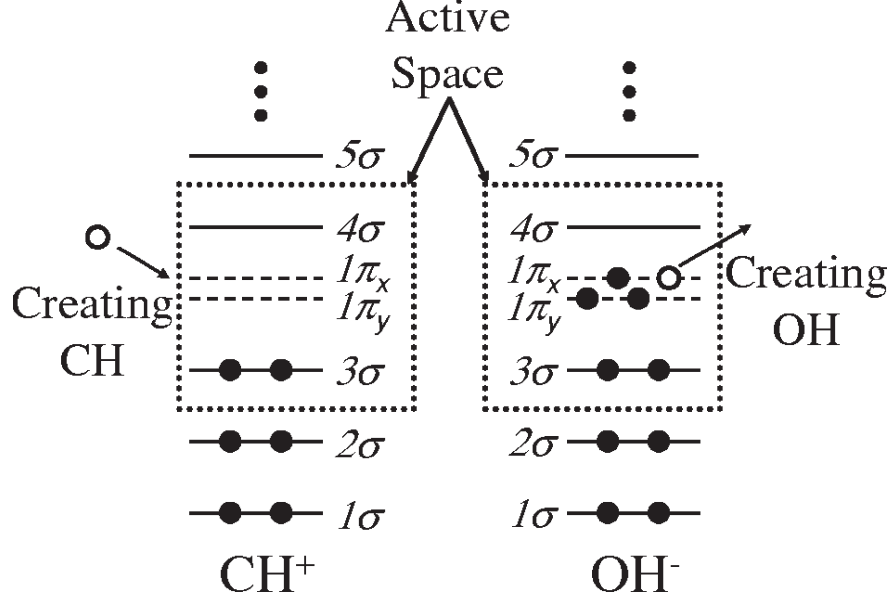


Figure 3.1: Orbital levels of the  $\text{CH}^+$  and  $\text{OH}^-$  ions and a schematic representation of the electron attachment and ionization processes that lead to the formation of the CH and OH radicals from the  $\text{CH}^+$  and  $\text{OH}^-$  reference closed-shell systems. Valence shells of  $\text{CH}^+$  and  $\text{OH}^-$  that play a dominant role in the relevant electron attachment and ionization processes and that are used in the active-space EA- and IP-EOMCC calculations are emphasized with the help of dotted frames.

of the active-space IP-EOMCCSD( $3h-2p$ ){ $N_o$ } approach are obtained by diagonalizing the similarity-transformed Hamiltonian of CCSD, Eq. (3.13), in the subspace of the  $(N - 1)$ -electron Hilbert space spanned by the  $|\Phi_i\rangle$ ,  $|\Phi_{ij}^b\rangle$ , and  $|\Phi_{\mathbf{I}jk}^{bc}\rangle$  determinants. This reduces the iterative  $n_o^3 n_u^4$  steps of the parent IP-EOMCCSD( $3h-2p$ ) approach to the  $N_o n_o^2 n_u^4$  steps, which are a small prefactor  $N_o$  ( $< n_o$ ) times the cost of the CCSD calculations, enabling one to perform the IP-EOMCCSD( $3h-2p$ )-level calculations for larger systems and basis sets than the full IP-EOMCCSD( $3h-2p$ ) scheme would allow.



## 3.2 Geometries and adiabatic excitation energies of the low-lying valence states of CNC, C<sub>2</sub>N, N<sub>3</sub>, and NCO

### 3.2.1 Computational details

To examine the performance of the full and active-space EA- and IP-EOMCC approaches with up to  $3p-2h/3h-2p$  excitations in the calculations of the ground and low-lying excited states of the CNC, C<sub>2</sub>N, NCO, and N<sub>3</sub> molecules and to answer the questions about the role of the method and basis set employed in determining the relevant nuclear geometries and adiabatic excitation energies, we performed the EA-EOMCCSD( $2p-1h$ ), active-space EA-EOMCCSD( $3p-2h$ ){4}, and full EA-EOMCCSD( $3p-2h$ ) calculations for the CNC and C<sub>2</sub>N molecules using the CCSD ground states of the CNC<sup>+</sup> and C<sub>2</sub>N<sup>+</sup> closed-shell cations to provide the reference wave functions and the IP-EOMCCSD( $2h-1p$ ), active-space IP-EOMCCSD( $3h-2p$ ){2}, and full IP-EOMCCSD( $3h-2p$ ) calculations for the NCO and N<sub>3</sub> molecules using the CCSD ground states of the NCO<sup>-</sup> and N<sub>3</sub><sup>-</sup> closed-shell anions to provide the reference wave functions. In addition to the previously employed [52, 179] DZP basis [185], we used the cc-pVDZ (all six methods), cc-pVTZ (all six methods), and cc-pVQZ (the EA-EOMCCSD( $2p-1h$ ), EA-EOMCCSD( $3p-2h$ ){4}, IP-EOMCCSD( $2h-1p$ ), and IP-EOMCCSD( $3h-2p$ ){2} approaches) basis sets [186]. The full EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ) calculations using the largest cc-pVQZ basis set could not be performed due to the large computer costs of the relevant numerical geometry optimizations (the analytic gradients of the EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ) approaches are not yet available). The active orbital spaces for the EA-EOMCCSD( $3p-2h$ ){4} calculations

for CNC and C<sub>2</sub>N consisted of the two lowest-energy pairs of unoccupied  $\pi$  molecular orbitals of CNC<sup>+</sup> and C<sub>2</sub>N<sup>+</sup>, respectively. The active orbital spaces for the IP-EOMCCSD( $3h-2p$ ){2} calculations for NCO and N<sub>3</sub> consisted of the highest-energy pair of occupied orbitals of NCO<sup>-</sup> and N<sub>3</sub><sup>-</sup>, respectively.

Unlike in the earlier work [52, 179], where the nuclear geometries of the ground and excited states of the CNC, C<sub>2</sub>N, NCO and N<sub>3</sub> species were optimized using only one method (SAC-CI-SDT-*R*/PS [187–189]) and one small basis set (DZP), in each molecular case and for each electronic state and basis set considered in the present work, the nuclear geometries were optimized at the same level of the EA/IP-EOMCC theory and with the same basis set as those used to evaluate the corresponding total and adiabatic excitation energies. The geometry optimizations using the cc-pVxZ ( $x = \text{D, T, Q}$ ) basis sets were constrained to linear geometries, since the analogous unconstrained optimizations using the DZP basis set and bent initial structures showed that the optimum geometries of all of the calculated states of CNC, C<sub>2</sub>N, NCO and N<sub>3</sub> are linear. The unconstrained optimizations with the DZP basis set demonstrated that we can assume the  $D_{\infty h}$  (in practice,  $D_{2h}$ ) symmetry for each of the calculated states of CNC and N<sub>3</sub>, and that we can use the  $C_{\infty v}$  (in practice,  $C_{2v}$ ) symmetry in the geometry optimizations for C<sub>2</sub>N and NCO. In all post-RHF (CCSD and EA/IP-EOMCC) calculations, the lowest-energy core orbitals correlating with the 1s orbitals of the C and N atoms were kept frozen and the spherical components of the  $d$ ,  $f$ , and  $g$  functions were employed throughout. All EA/IP-EOMCC calculations reported in this study were performed using the EA-EOMCCSD( $2p-1h$ ), IP-EOMCCSD( $2h-1p$ ), and full and active-space EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ) routines described

in Ref. [159] and incorporated by the Piecuch group in the GAMESS package [190, 191], which were interfaced with the numerical derivative GAMESS codes by Dr. Michael W. Schmidt.

In addition to the results of the finite basis set calculations using the cc-pV $x$ Z ( $x = \text{D, T, Q}$ ) basis sets, we extrapolated the total and excitation energies obtained in the EA- and IP-EOMCC calculations for the CNC, C<sub>2</sub>N, NCO and N<sub>3</sub> molecules to the complete basis set (CBS) limit. We had to limit our CBS extrapolations to the EA-EOMCCSD( $2p-1h$ ), IP-EOMCCSD( $2h-1p$ ), active-space EA-EOMCCSD( $3p-2h$ ){4}, and active space IP-EOMCCSD( $3h-2p$ ){2} calculations, since we were unable to carry out the full EA-EOMCCSD( $3p-2h$ ) and IP-EOMCCSD( $3h-2p$ ) computations including geometry optimizations using the cc-pVQZ basis set. In all of the remaining cases, we had access to the complete data sets corresponding to the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets, enabling the reasonably meaningful CBS extrapolations. In analogy to the earlier study of the low-lying states of methylene [192] and to verify the numerical stability of our CBS extrapolations, two different extrapolation schemes, referred to as the CBS-A and CBS-B approaches, were utilized in this work.

In the CBS-A scheme, we first determined the CBS limit of the ground-state CCSD correlation energy of the closed-shell  $N$ -electron reference system relevant to the EA/IP-EOMCC calculations for the  $(N \pm 1)$ -electron target species using the well-known extrapolation formula [193]

$$\Delta E(x) = \Delta E_{\infty} + Ax^{-3}, \quad (3.17)$$

and the cc-pVTZ and cc-pVQZ data. Here,  $\Delta E(x)$  is the CCSD correlation energy obtained

with the cc-pV $x$ Z basis, where  $x$  represents the cardinal number of the basis set ( $x = 3$  for cc-pVTZ and  $x = 4$  for cc-pVQZ), and  $\Delta E_\infty$  is the CCSD correlation energy in the CBS limit. The resulting extrapolated correlation energy  $\Delta E_\infty$  was then added to the RHF/cc-pV6Z energy of the  $N$ -electron reference system, computed at the optimized geometry of the state of interest resulting from the appropriate EA- or IP-EOMCC/cc-pVQZ calculations. In doing so, we used the well-known fact that the RHF energies converge exponentially with the basis set and it is usually better to determine the CBS limit of the RHF energy by performing the calculations with the very large correlation consistent basis set, such as cc-pV6Z, if such calculations are affordable (we verified the level of basis-set convergence of the RHF/cc-pV6Z calculations by comparing the RHF/cc-pV5Z and RHF/cc-pV6Z data, obtaining the differences of about 0.2 millihartree in all of the examined cases). Once the CBS values of the RHF and CCSD energies of the  $N$ -electron reference system were determined, we computed the desired CBS limits of the ground- and excited-state energies of the  $(N \pm 1)$ -electron target species corresponding to the EA- or IP-EOMCC calculations of interest using the formula

$$\begin{aligned}
E_{\mu,\infty}^{\text{EA/IP-EOMCC}}(N \pm 1) = & E_{6Z}^{\text{RHF}}(N) + \Delta E_{0,\infty}^{\text{CCSD}}(N) \\
& + E_{\mu,\text{QZ}}^{\text{EA/IP-EOMCC}}(N \pm 1) - E_{0,\text{QZ}}^{\text{CCSD}}(N),
\end{aligned} \tag{3.18}$$

where  $E_{\mu,\infty}^{\text{EA/IP-EOMCC}}(N \pm 1)$  is the final extrapolated energy of the  $(N \pm 1)$ -electron excited state  $|\Psi_\mu^{(N \pm 1)}\rangle$ ,  $E_{6Z}^{\text{RHF}}(N)$  is the ground-state RHF energy of the closed-shell  $N$ -electron reference system obtained with the cc-pV6Z basis set,  $\Delta E_{0,\infty}^{\text{CCSD}}(N)$  is the extrapolated CCSD correlation energy of the  $N$ -electron reference system obtained using Eq. (3.17),  $E_{\mu,\text{QZ}}^{\text{EA/IP-EOMCC}}(N \pm 1)$  is the total EA/IP-EOMCC energy of the  $(N \pm 1)$ -electron excited

state  $|\Psi_\mu^{(N\pm 1)}\rangle$  obtained with the cc-pVQZ basis set, and  $E_{0,\text{QZ}}^{\text{CCSD}}(N)$  is the total CCSD energy of the  $N$ -electron reference system obtained using the cc-pVQZ basis set (note that the difference of the last two energies in Eq. (3.18) is the electron-attached or ionized energy  $\omega_\mu^{(N\pm 1)}$ , Eq. (3.9), resulting from the relevant EA/IP-EOMCC diagonalization of  $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$ , obtained with the cc-pVQZ basis). This method of estimating the CBS values of the total electronic energies of the ground and excited states of the CNC, C<sub>2</sub>N, NCO and N<sub>3</sub> radicals is based on the assumption that the electron-attachment or ionization energies  $\omega_\mu^{(N\pm 1)}$ , Eq. (3.9), obtained with the cc-pVQZ basis set are essentially converged with respect to the basis set, so all one has to do is obtain the CBS limit of the CCSD ground-state energy of the  $N$ -electron reference system and add the cc-pVQZ values of the electron-attachment or ionization energies to estimate the CBS energies of the ground and excited states of the corresponding  $(N \pm 1)$ -electron target species. The validity of this assumption is discussed in what follows.

In the second basis set extrapolation method, referred to as the CBS-B approach, the total CBS energy of each  $(N \pm 1)$ -electron target state of interest was directly extrapolated using the formula [194]

$$E(x) = E_\infty + Be^{-(x-1)} + Ce^{-(x-1)^2} \quad (3.19)$$

and the cc-pVDZ, cc-pVTZ, and cc-pVQZ data. As in Eq. (3.17), the  $x$  variable number entering Eq. (3.19) is the cardinal number of the cc-pV $x$ Z basis set ( $x = 2$  for cc-pVDZ,  $x = 3$  for cc-pVTZ, and  $x = 4$  for cc-pVQZ),  $E(x)$  is the total EA/IP-EOMCC energy computed with the cc-pV $x$ Z basis set, and  $E_\infty$  is the desired CBS limit of the total EA/IP-EOMCC

energy for a given electronic state of the  $(N \pm 1)$ -electron species. The difference between the CBS-A and CBS-B extrapolation schemes lies in the fact that the latter scheme extrapolates the total EA- or IP-EOMCC energy of each electronic state of the  $(N \pm 1)$ -electron target species separately, using Eq. (3.19), whereas the former approach extrapolates the ground-state energy of the  $N$ -electron reference system only using Eq. (3.17), while making an assumption that the electron-attachment and ionization energies resulting from the EA- and IP-EOMCC calculations converge faster with the basis set than the total energies of the  $(N \pm 1)$ -electron target species, as reflected in Eq. (3.18). A comparison of both extrapolation schemes will tell us how accurate this assumption is.

### 3.2.2 Results

The results of our EA- and IP-EOMCC calculations, along with the available experimental data [195–197], are reported in Tables 3.1–3.6. The EA-EOMCCSD( $2p-1h$ ), EA-EOMCCSD( $3p-2h$ ){4}, and full EA-EOMCCSD( $3p-2h$ ) results for the CNC and C<sub>2</sub>N molecules are reported in Tables 3.1 and 3.2 for the total and adiabatic excitation energies, and 3.3 for the geometries. The IP-EOMCCSD( $2h-1p$ ), IP-EOMCCSD( $3h-2p$ ){2}, and full IP-EOMCCSD( $3h-2p$ ) results for the NCO and N<sub>3</sub> molecules are reported in Tables 3.4 and 3.5 for the total and adiabatic excitation energies, and 3.6 for the geometries. Our discussion is divided into two subsections. Subsection 3.2.2.1 examines the EA-EOMCC results for CNC and C<sub>2</sub>N. The IP-EOMCC calculations for NCO and N<sub>3</sub> are discussed in Subsection 3.2.2.2.

### 3.2.2.1 EA-EOMCC results for CNC and C<sub>2</sub>N

Let us begin our discussion with the EA-EOMCC calculations with the CNC molecule (see Tables 3.1 and 3.3). For CNC, the EA-EOMCC optimizations employing the DZP basis set produced results that deviate from the previously reported [179] SAC-CI-SDT-*R*/PS optimized geometries and previously calculated [52, 179] EA-EOMCC adiabatic excitation energies using the SAC-CI-SDT-*R*/PS geometries, all obtained with the same DZP basis, by 0.001-0.009 Å and 0.002-0.003 eV, respectively, for all states and all methods considered in this work. Seeing that our present EA-EOMCC optimizations employing the DZP basis set reproduced the analogous results of the SAC-CI-SDT-*R*/PS geometry optimizations and the EA-EOMCC excitation energies at the SAC-CI-SDT-*R*/PS geometries reported in Ref. [179], we proceeded to study the effect of the use of the correlation-consistent basis sets of the cc-pVxZ quality on the calculated excitation energies and geometries. As mentioned in the previous section, in all of the EA-EOMCC calculations for CNC using the cc-pVxZ basis sets, reported in this study, we used the  $D_{2h}$  Abelian symmetry. Thus, the ground  $X\ ^2\Pi_g$  state of CNC was calculated as the lowest-energy  $^2B_{2g}$  state, although we could also use the lowest-energy  $^2B_{3g}$  state to represent the doubly (if we ignore spin) degenerate  $X\ ^2\Pi_g$  state, obtaining exactly the same results. The  $A\ ^2\Delta_u$  state was calculated as the lowest-energy  $^2A_u$  state, although we could obtain the equivalent energies and geometries of this doubly (again, ignoring spin) degenerate state by considering the lowest-energy  $^2B_{1u}$  state (which was useful to avoid confusion with the spatially non-degenerate  $B\ ^2\Sigma_u^+$  state). The  $B\ ^2\Sigma_u^+$  state was calculated as the second  $^2B_{1u}$  state. Unless otherwise specified, the discussion below focuses on the EA-EOMCC calculations using the cc-pVxZ ( $x = D, T, Q$ ) basis sets.

Table 3.1: Total and adiabatic excitation energies for the ground and low-lying excited states of CNC, as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ ( $x = D, T, Q$ ) basis sets and extrapolating to the CBS limit.

| Method                   | Basis | Total Energy (hartree) |                  |                    | Adiabatic Excitation Energy (eV) |                                  |
|--------------------------|-------|------------------------|------------------|--------------------|----------------------------------|----------------------------------|
|                          |       | $X \ ^2\Pi_g$          | $A \ ^2\Delta_u$ | $B \ ^2\Sigma_u^+$ | $A \ ^2\Delta_u - X \ ^2\Pi_g$   | $B \ ^2\Sigma_u^+ - X \ ^2\Pi_g$ |
| EA-EOMCCSD( $2p-1h$ )    | DZP   | -130.406718            | -130.141822      | -130.125873        | 7.208                            | 7.642                            |
|                          | x=D   | -130.402813            | -130.136443      | -130.120048        | 7.248                            | 7.694                            |
|                          | x=T   | -130.502669            | -130.220645      | -130.204320        | 7.674                            | 8.118                            |
|                          | x=Q   | -130.534268            | -130.248264      | -130.232033        | 7.783                            | 8.224                            |
|                          | CBS-A | -130.551020            | -130.264878      | -130.248586        | 7.786                            | 8.230                            |
|                          | CBS-B | -130.552172            | -130.264020      | -130.247849        | 7.841                            | 8.281                            |
| EA-EOMCCSD( $3p-2h$ )    | DZP   | -130.411686            | -130.260720      | -130.238177        | 4.108                            | 4.721                            |
|                          | x=D   | -130.408191            | -130.257611      | -130.234329        | 4.097                            | 4.731                            |
|                          | x=T   | -130.510334            | -130.358548      | -130.335201        | 4.130                            | 4.766                            |
| EA-EOMCCSD( $3p-2h$ ){4} | DZP   | -130.409784            | -130.259560      | -130.236779        | 4.088                            | 4.708                            |
|                          | x=D   | -130.406511            | -130.256819      | -130.233332        | 4.073                            | 4.712                            |
|                          | x=T   | -130.507154            | -130.356797      | -130.333074        | 4.091                            | 4.737                            |
|                          | x=Q   | -130.538435            | -130.388104      | -130.364472        | 4.091                            | 4.734                            |
|                          | CBS-A | -130.554997            | -130.404664      | -130.380953        | 4.091                            | 4.736                            |
|                          | CBS-B | -130.556095            | -130.405806      | -130.382242        | 4.090                            | 4.731                            |
| Experiment <sup>a</sup>  |       |                        |                  |                    | 3.761                            | 4.315                            |

<sup>a</sup> Taken from Refs. [195, 196].



The ground  $X\ ^2\Pi_g$  state of CNC is dominated by  $1p$  excitations out of the ground state of the closed-shell reference  $\text{CNC}^+$  ion, which helps the low-order EA-EOMCC calculations for this state. However, the  $A\ ^2\Delta_u$  and  $B\ ^2\Sigma_u^+$  excited states of CNC exhibit a significant two-electron excitation character relative to the  $X\ ^2\Pi_g$  state, resulting in the large  $2p-1h$  contributions in the corresponding wave functions relative to the ground state of  $\text{CNC}^+$  which reflect on the more complex multireference nature of these two states. As a result and as shown in Table 3.1, the basic EA-EOMCCSD( $2p-1h$ ) optimizations produced adiabatic excitation energies that deviate from the experimental values by 3.379-4.022 eV for the  $A\ ^2\Delta_u$  and  $B\ ^2\Sigma_u^+$  states, demonstrating the same characteristically large errors compared to experiment that we typically see when the EA-EOMCCSD( $2p-1h$ ) approach is applied to the excited states of radicals dominated by two-electron transitions [157–160]. The full EA-EOMCCSD( $3p-2h$ ) method improves these poor results, reducing the deviations from experiment to about 0.336-0.451 eV for both the  $A\ ^2\Delta_u$  and  $B\ ^2\Sigma_u^+$  states, when the cc-pVDZ and cc-pVTZ basis sets are employed. The reason for this considerable improvement in the data over the EA-EOMCCSD( $2p-1h$ ) method is the explicit inclusion of the  $3p-2h$  terms in the  $R_\mu^{(N+1)}$  operator in the EA-EOMCCSD( $3p-2h$ ) calculations.

As explained in the previous section, the inclusion of all  $3p-2h$  components in the  $R_\mu^{(N+1)}$  operator is computationally demanding, particularly when one has to perform numerical gradient optimizations, such as those reported in this work. Thus, it is of great significance to note that the active-space EA-EOMCCSD( $3p-2h$ ){4} optimizations using the cc-pVDZ and cc-pVTZ basis sets, with only four unoccupied orbitals in the active space, which are only a few times more expensive than the corresponding ground-state CCSD calculations

and which require a small fraction of the CPU time, disk, and memory when compared to the parent EA-EOMCCSD( $3p-2h$ ) calculations, reproduce the full EA-EOMCCSD( $3p-2h$ ) optimization results to within 0.019-0.039 eV for the adiabatic excitation energies and 0.8-3.2 millihartree for the total energies of the ground and excited states of CNC examined in this study (see Table 3.1). The deviations of the EA-EOMCCSD( $3p-2h$ ){4} results from experiment are 0.312-0.422 eV for all three correlation consistent basis sets used in the EA-EOMCCSD( $3p-2h$ ){4} optimizations. It is interesting and somewhat surprising to note the deviations from experiment increase slightly with the size of the cc-pV $x$ Z basis set for all three EA-EOMCC methods exploited in this work. To help understand this behavior, we extrapolated our EA-EOMCCSD( $2p-1h$ ) and EA-EOMCCSD( $3p-2h$ ){4} results to the CBS limit.

Examining the total energies of the ground and excited states of CNC shown in Table 3.1, it is clear that they are converging with the basis set in a systematic manner. The CBS-A extrapolation scheme is based on the simplifying assumption that the electron-attachment (or, in the IP case, ionization) energies are reasonably well converged with the basis set, when the cc-pVQZ basis set is employed. The data in Table 3.1 show that this is indeed a valid assumption, as the EA-EOMCCSD( $2p-1h$ ) and EA-EOMCCSD( $3p-2h$ ){4} excitation energies do not significantly change when moving from the cc-pVTZ to cc-pVQZ basis, the largest change being 0.109 eV for the less accurate EA-EOMCCSD( $2p-1h$ ) and only 0.003 eV for the active-space EA-EOMCCSD( $3p-2h$ ){4} approach. Moreover, the CBS extrapolations resulting from the CBS-A and CBS-B schemes produce results that are in good agreement, especially for the higher-order EA-EOMCCSD( $3p-2h$ ){4} method, where the differences in

total energies do not exceed 1.3 millihartree, regardless of the electronic state of CNC considered (we recall that the CBS-B scheme extrapolates the total energy of each state separately, without any simplifying assumptions). The CBS limits of the excitation energies resulting from the CBS-A and CBS-B extrapolations obtained with the EA-EOMCCSD( $3p-2h$ ){4} approach are essentially identical, deviating by 0.001 eV and 0.005 eV for the  $A\ ^2\Delta_u$  and  $B\ ^2\Sigma_u^+$  states, respectively. Thus, we can safely conclude that the CBS EA-EOMCCSD( $3p-2h$ ){4} results are stable to approximately 1 millihartree for the total energies and 0.005 eV for the adiabatic excitation energies, and can be regarded as converged with the basis set.

The CBS-A and CBS-B excitation energies obtained with the EA-EOMCCSD( $3p-2h$ ){4} approach deviate from the available experimental values by no more than 0.330 eV for the  $A\ ^2\Delta_u$  state and no more than 0.421 eV for the  $B\ ^2\Sigma_u^+$  state. In Ref. [198], the adiabatic excitation energies of CNC were calculated using the CASPT2 approach and the CASSCF-based MRCI (multireference CI) calculations. The CASPT2 results reported in Ref. [198] deviate from experiment by 0.297 eV for the  $A\ ^2\Delta_u$  state and 0.009 eV for the  $B\ ^2\Sigma_u^+$  state. The latter value must be a result of the fortuitous cancellation of errors, since the CASSCF-based MRCI calculations reported in the same work give the deviations from experiment which are 0.253 eV for the  $A\ ^2\Delta_u$  state and 0.328 eV for the  $B\ ^2\Sigma_u^+$  state. We can see that our computationally much less demanding and much easier to use, largely single-reference, active-space EA-EOMCCSD( $3p-2h$ ) calculations can account for the multireference character of the excited states of open-shell species, such as the  $A\ ^2\Delta_u$  and  $B\ ^2\Sigma_u^+$  states of CNC, without having to sacrifice the accuracies of genuine high-level multireference calculations. Our converged EA-EOMCCSD( $3p-2h$ ){4} data and their EA-EOMCCSD( $3p-2h$ )/cc-pVTZ

analogs indicate one of the following two things: (i) the higher-order (e.g.  $4p-3h$ ) excitations are needed in the EA-EOMCC calculations to close the gap between the experimental and theoretical data or (ii) the experiments that we use here to assess the quality of our calculations have some unaccounted errors that may have to be reexamined.

For the  $C_2N$  molecule (Tables 3.2 and 3.3), the EA-EOMCC optimizations employing the DZP basis carried out in this study produced results that are very similar to the previously reported [52,179] EA-EOMCC adiabatic excitation energies calculated at the SAC-CI-SDT- $R/PS$  optimized geometries, all obtained with the same DZP basis as that used here. For example, the deviations between the adiabatic excitation energies obtained using the present EA-EOMCC/DZP optimizations and the analogous excitation energies reported in Refs. [52,179], which used the SAC-CI-SDT- $R/PS$  geometries, are 0.087-0.155 eV for the EA-EOMCCSD( $2p-1h$ ) case and 0.000-0.003 eV for the full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} cases, for all three excited states of  $C_2N$  examined in this study. Again, seeing that our present EA-EOMCC optimizations employing the DZP basis set reproduced the analogous results calculated at the geometries obtained with the SAC-CI-SDT- $R/PS$  approach, reported in Refs. [52,179], we proceeded to study the effect of the use of the correlation-consistent basis sets of the cc-pV $x$ Z quality on the calculated excitation energies and geometries. As mentioned in Section 3.2.1, in all of the EA-EOMCC calculations for  $C_2N$  using the cc-pV $x$ Z basis sets, discussed in the present paper, we used the  $C_{2v}$  Abelian symmetry. Thus, the ground  $X\ ^2\Pi$  state of  $C_2N$  was calculated as the lowest-energy  $^2B_1$  state, although we could also use the lowest-energy  $^2B_2$  state, obtaining identical results. The doubly (ignoring spin) degenerate  $A\ ^2\Delta$  state was calculated as the

lowest-energy  ${}^2A_1$  or  ${}^2A_2$  state (we carried out both calculations to verify the correctness of our results and to avoid confusion with the spatially non-degenerate  $B\ {}^2\Sigma^-$  state), the  $B\ {}^2\Sigma^-$  state was calculated as the second  ${}^2A_2$  state, and the  $C\ {}^2\Sigma^+$  state was obtained as the second  ${}^2A_1$  state. As in the CNC case, our discussion of the results obtained for  $C_2N$  focuses on the EA-EOMCC calculations using the cc-pV $x$ Z ( $x = D, T, Q$ ) basis sets.

In analogy to CNC, the ground  $X\ {}^2\Pi$  state of  $C_2N$  is dominated by  $1p$  excitations out of the ground state of the closed-shell reference  $C_2N^+$  ion, but the low-lying  $A\ {}^2\Delta$ ,  $B\ {}^2\Sigma^-$ , and  $C\ {}^2\Sigma^+$  excited states have significant  $2p-1h$  contributions demonstrating the rather complex multireference nature of their corresponding wave functions. The  $B\ {}^2\Sigma^-$  state also has non-negligible  $3p-2h$  contributions, which make this state extremely difficult to describe. All of this causes major problems in the EA-EOMCCSD( $2p-1h$ ) calculations. As shown in Table 3.2, even with a large cc-pVQZ basis set, the EA-EOMCCSD( $2p-1h$ ) method incorrectly orders the excited states of  $C_2N$ , describing the  $C\ {}^2\Sigma^+$  state as being lower in energy than the  $B\ {}^2\Sigma^-$  state. The errors in the EA-EOMCCSD( $2p-1h$ ) results for the adiabatic excitation energies of  $C_2N$  relative to experiment are huge. Indeed, our geometry optimizations using the EA-EOMCCSD( $2p-1h$ ) approach produce errors in the calculated adiabatic excitation energies of  $C_2N$  relative to experiment of 3.507-3.969 eV for the  $A\ {}^2\Delta$  state, 4.956-5.511 eV for the  $B\ {}^2\Sigma^-$  state, and 3.422-3.836 eV for the  $C\ {}^2\Sigma^+$  state. As with CNC, the full inclusion of the  $3p-2h$  components in the electron attaching operator  $R_\mu^{(N+1)}$  significantly improves the adiabatic excitation energies relative to the disastrous EA-EOMCCSD( $2p-1h$ ) results, reducing the errors relative to experiment to at most 0.418 eV for the  $A\ {}^2\Delta$  state, at most 0.915 eV for the  $B\ {}^2\Sigma^-$  state, and at most 0.524 eV for

the  $C\ ^2\Sigma^+$  state when the cc-pVDZ and cc-pVTZ basis sets are employed, but the full EA-EOMCCSD( $3p-2h$ ) are computationally demanding, particularly when larger basis sets have to be examined. Thus, it is important to examine how well the considerably less expensive active-space EA-EOMCCSD( $3p-2h$ ) approach works for the low-lying excited states of  $C_2N$ , when the cc-pV $x$ Z basis sets are employed.

Table 3.2: Total and adiabatic excitation energies for the ground and low-lying excited states of C<sub>2</sub>N obtained with the various EA-EOMCC approaches using the DZP [4s2p1d] and cc-pV $x$ Z ( $x = D, T, Q$ ) basis sets and extrapolating to the CBS limit.

| Method                   | Basis | Total Energy (hartree) |             |               |               | Adiabatic Excitation<br>Energy (eV) |                           |                           |
|--------------------------|-------|------------------------|-------------|---------------|---------------|-------------------------------------|---------------------------|---------------------------|
|                          |       | $X^2\Pi$               | $A^2\Delta$ | $B^2\Sigma^-$ | $C^2\Sigma^+$ | $A^2\Delta^-$<br>$X^2\Pi$           | $B^2\Sigma^-$<br>$X^2\Pi$ | $C^2\Sigma^+$<br>$X^2\Pi$ |
|                          |       |                        |             |               |               |                                     |                           |                           |
| EA-EOMCCSD( $2p-1h$ )    | DZP   | -130.400501            | -130.176452 | -130.117500   | -130.156651   | 6.097                               | 7.701                     | 6.635                     |
|                          | x=D   | -130.400086            | -130.174345 | -130.115824   | -130.152828   | 6.143                               | 7.735                     | 6.728                     |
|                          | x=T   | -130.499176            | -130.259914 | -130.198940   | -130.240134   | 6.511                               | 8.170                     | 7.049                     |
|                          | x=Q   | -130.530280            | -130.287558 | -130.225643   | -130.267832   | 6.605                               | 8.290                     | 7.142                     |
|                          | CBS-A | -130.546521            | -130.303831 | -130.241794   | -130.283903   | 6.604                               | 8.292                     | 7.146                     |
|                          | CBS-B | -130.548409            | -130.303738 | -130.241271   | -130.284090   | 6.658                               | 8.358                     | 7.193                     |
| EA-EOMCCSD( $3p-2h$ )    | DZP   | -130.405260            | -130.292989 | -130.270231   | -130.265181   | 3.055                               | 3.674                     | 3.812                     |
|                          | x=D   | -130.404842            | -130.292610 | -130.270337   | -130.264086   | 3.054                               | 3.660                     | 3.830                     |
|                          | x=T   | -130.506456            | -130.394642 | -130.370688   | -130.366299   | 3.043                               | 3.694                     | 3.814                     |
| EA-EOMCCSD( $3p-2h$ ){4} | DZP   | -130.403651            | -130.292385 | -130.269696   | -130.264361   | 3.028                               | 3.645                     | 3.791                     |
|                          | x=D   | -130.403260            | -130.292089 | -130.269870   | -130.263371   | 3.025                               | 3.630                     | 3.807                     |
|                          | x=T   | -130.503555            | -130.393547 | -130.369756   | -130.364858   | 2.993                               | 3.641                     | 3.774                     |
|                          | x=Q   | -130.533052            | -130.423692 | -130.399922   | -130.394864   | 2.976                               | 3.623                     | 3.760                     |
|                          | CBS-A | -130.543559            | -130.433071 | -130.408605   | -130.404035   | 3.007                               | 3.672                     | 3.797                     |
|                          | CBS-B | -130.549517            | -130.440555 | -130.416854   | -130.411633   | 2.965                               | 3.610                     | 3.752                     |
| Experiment <sup>a</sup>  |       |                        |             |               |               | 2.636                               | 2.779                     | 3.306                     |

<sup>a</sup> Taken from Refs. [195, 197].

As shown in Table 3.2, the results of the active-space EA-EOMCCSD( $3p-2h$ ){4} calculations are almost identical to those obtained with the parent EA-EOMCCSD( $3p-2h$ ) approach. The adiabatic excitation energies obtained with the full and active space EA-EOMCCSD( $3p-2h$ ) methods, where the latter approach uses only four unoccupied orbitals in the active space, calculated using the cc-pVDZ and cc-pVTZ basis sets, differ by 0.023-0.053 eV for all states of C<sub>2</sub>N examined in this work. The total energies obtained in the full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} calculations employing the cc-pVDZ and cc-pVTZ basis sets differ by 1.6-2.9 millihartree for the  $X\ ^2\Pi$  state, 0.5-1.1 millihartree for the  $A\ ^2\Delta$  state, 0.5-0.9 millihartree for the  $B\ ^2\Sigma^-$  state, and 0.7-1.4 millihartree for the  $C\ ^2\Sigma^+$  state. We would like to emphasize once again that the active-space EA-EOMCCSD( $3p-2h$ ){4} calculations provide results with comparable accuracy to its more expensive EA-EOMCCSD( $3p-2h$ ) parent method at a small fraction of the computational cost and with an effort which is on the same order as that characterizing the standard CCSD calculations. Comparing the EA-EOMCCSD( $3p-2h$ ){4} results with experiment, we see that the adiabatic excitation energies resulting from the EA-EOMCCSD( $3p-2h$ ){4} calculations using the cc-pVQZ basis set differ from the available experimental data by 0.340 eV, 0.844 eV, and 0.454 eV for the  $A\ ^2\Delta$ ,  $B\ ^2\Sigma^-$ , and  $C\ ^2\Sigma^+$  states, respectively, which is a huge error reduction when compared to the corresponding EA-EOMCCSD( $2p-1h$ )/cc-pVQZ calculations that give the 3.969 eV, 5.511 eV, and 3.836 eV errors for the same three states, in addition to wrong state ordering. The full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} calculations produce the correct state ordering and relatively small errors for the  $A\ ^2\Delta$  and  $C\ ^2\Sigma^+$  states, but the discrepancy between the full and active-space



EA-EOMCCSD( $3p-2h$ ) results on the one hand and experiment on the other hand for the  $B\ ^2\Sigma^-$  state, on the order of 0.9 eV independent of the basis set, is a problem that needs to be addressed.

The larger deviations with experiment observed in the full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} calculations for the  $B\ ^2\Sigma^-$  state, which do not seem to be decreasing with the basis set and careful geometry optimizations performed in this work, must be related to the presence of the non-negligible  $3p-2h$  contributions in the  $B\ ^2\Sigma^-$  wave function, which indicate a highly multireference character of this state that the EA-EOMCC methods used in the present study cannot capture without incorporating higher-than- $3p-2h$  contributions in the EA-EOMCC considerations. As explained in Ref. [160], the presence of significant  $3p-2h$  contributions in the wave function requires an explicit consideration of the  $4p-3h$  and, perhaps, higher-than- $4p-3h$  components of the  $R_\mu^{(N+1)}$  operator in the EA-EOMCC calculations, neglected at the EA-EOMCCSD( $3p-2h$ ) level. The highly multireference character of the  $B\ ^2\Sigma^-$  state becomes clear when we examine the CASPT2 and CASSCF-based MRCI calculations reported in Ref. [198]. These calculations are in reasonable agreement with the results of our full and active-space EA-EOMCCSD( $3p-2h$ ) calculations for the  $A\ ^2\Delta$  and  $C\ ^2\Sigma^+$  states, producing the 0.238 eV error for the  $A\ ^2\Delta$  state and 0.219 eV error for the  $C\ ^2\Sigma^+$  state when the CASPT2 approach is employed, but the CASPT2 and MRCI results obtained in Ref. [198] for the  $B\ ^2\Sigma^-$  state are considerably more accurate than those obtained here with the full and active-space EA-EOMCCSD( $3p-2h$ ) theory levels. Indeed, the CASPT2 and MRCI calculations for the  $B\ ^2\Sigma^-$  state reported in Ref. [198] give the 0.225 eV and 0.250 eV errors, respectively, relative to experiment, as

opposed to  $\sim 0.9$  eV obtained with the full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} methods. Interestingly, the MRCI approach improves the CASPT2 results for the  $A\ ^2\Delta$  and  $C\ ^2\Sigma^+$  states as well, reducing the 0.238 eV and 0.219 eV errors obtained in the CASPT2 calculations to 0.060 eV and 0.058 eV, respectively [198], which suggests that the incorporation of the  $4p-3h$  and, perhaps, some other higher-order excitations in the EA-EOMCC calculations may be necessary to further improve the description of all three excited states of  $C_2N$  examined in this work. Since the calculations reported in the present paper exclude the possibility that the basis set or geometry optimizations may help the EA-EOMCCSD( $3p-2h$ ) results, the next logical step would be to examine the role of  $4p-3h$  excitations in the EA-EOMCC calculations. One may also have to examine if the use of the full CCSDT or active-space CCSDt approaches rather than the CCSD method in providing the ground-state wave function for the reference  $C_2N^+$  ion plays a role here. These will be the topics of future work.

In distinct juxtaposition to the CNC case, as larger basis sets are employed, the EA-EOMCCSD( $3p-2h$ ) and EA-EOMCCSD( $3p-2h$ ){4} adiabatic excitation energies for all three states of  $C_2N$  investigated in the present study become slightly smaller and slightly closer to the experimental values or do not change, although, as already pointed out, the significant errors remain even when the larger basis sets of the cc-pVTZ or cc-pVQZ quality are employed. Interestingly, the EA-EOMCCSD( $2p-1h$ ) results have deviations from experiment that seem to grow with the size of the basis set. It is, therefore, worthwhile to examine the CBS limits of our EA-EOMCC results for the  $C_2N$  system.

As shown in Table 3.2, the adiabatic excitation energies resulting from the EA-EOMCCSD

$(2p-1h)$  and EA-EOMCCSD( $3p-2h$ ){4} calculations with the cc-pVQZ basis set are reasonably well converged with the basis and, although the total energies of the individual electronic states of  $C_2N$  are not converged when the cc-pVQZ basis set is employed, they behave in a systematic manner as we go from the cc-pVDZ to cc-pVQZ basis sets, facilitating the CBS extrapolations. Indeed, when going from the cc-pVTZ to the cc-pVQZ basis sets, the changes in the EA-EOMCCSD( $3p-2h$ ){4} excitation energies are very small, at most 0.018 eV. The analogous changes in the EA-EOMCCSD( $2p-1h$ ) excitation energies are somewhat larger (at most 0.120 eV), but we can still view them as reasonably stable considering the complicated nature of the  $C_2N$  excited states that the EA-EOMCCSD( $2p-1h$ ) approach has significant problems with. Overall, the simplifying assumption of the CBS-A extrapolation scheme that one can treat the electron-attachment energies resulting from the EA-EOMCC calculations with the cc-pVQZ basis set as essentially converged values remains valid for  $C_2N$ , so we expect the CBS-A scheme to provide meaningful results. This can be verified by comparing the CBS-A and CBS-B extrapolations. Comparing the EA-EOMCCSD( $2p-1h$ ) CBS-A and CBS-B values, the total energies differ by 1.9 millihartree for the  $X^2\Pi$  state, 0.1 millihartree for the  $A^2\Delta$  state, 0.5 millihartree for the  $B^2\Sigma^-$  state, and 0.2 millihartree for the  $C^2\Sigma^+$  state. The adiabatic excitation energies resulting from the CBS-A and CBS-B extrapolations of the EA-EOMCCSD( $2p-1h$ ) data differ by 0.054 eV for the  $A^2\Delta$  state, 0.066 eV for the  $B^2\Sigma^-$  state, and 0.047 eV for the  $C^2\Sigma^+$  state. The CBS-A and CBS-B results for the EA-EOMCCSD( $3p-2h$ ){4} total energies differ by 6.0 millihartree for the  $X^2\Pi$  state, 7.5 millihartree for the  $A^2\Delta$  state, 8.2 millihartree for the  $B^2\Sigma^-$  state, and 7.6 millihartree for the  $C^2\Sigma^+$  state. The differences in the adiabatic excitation energies obtained with the two

CBS extrapolation schemes, as applied to the EA-EOMCCSD( $3p-2h$ ){4} data, are 0.042 eV, 0.062 eV, and 0.045 eV for the  $A\ ^2\Delta$ ,  $B\ ^2\Sigma^-$  and  $C\ ^2\Sigma^+$  states, respectively. We can conclude that our CBS EA-EOMCC results for the  $C_2N$  molecule are generally stable to within about 8 millihartree for the total energies and 0.060 eV for the adiabatic excitation energies. The deviations of the CBS-A extrapolated EA-EOMCCSD( $3p-2h$ ){4} results from experiment are 0.371 eV for the  $A\ ^2\Delta$  state, 0.893 eV for the  $B\ ^2\Sigma^-$  state, and 0.491 eV for the  $C\ ^2\Sigma^+$  state. The analogous CBS-B calculations employing the EA-EOMCCSD( $3p-2h$ ){4} data give the 0.329 eV, 0.831 eV, and 0.446 eV errors, respectively. These results indicate once again that higher-than- $3p-2h$  excitations and, perhaps, methods better than CCSD for the description of the ground state of the reference  $C_2N^+$  ion may have to be included in the EA-EOMCC calculations for the low-lying states of the  $C_2N$  molecule, particularly in the case of the  $B\ ^2\Sigma^-$  state.

Table 3.3: Comparison of the optimized equilibrium geometries for the low-lying states of CNC and C<sub>2</sub>N, as obtained with the EA-EOMCC and SAC-CI-SDT-*R*/PS approaches using the DZP[4*s*2*p*1*d*] and cc-pV*x*Z (*x* = D, T, Q) basis sets.

| Method                                 | Basis | CNC <sup>a</sup>                            |                                             |                                                          | C <sub>2</sub> N <sup>b</sup> |                         |                                      |                                      |
|----------------------------------------|-------|---------------------------------------------|---------------------------------------------|----------------------------------------------------------|-------------------------------|-------------------------|--------------------------------------|--------------------------------------|
|                                        |       | <i>X</i> <sup>2</sup> Π <sub><i>g</i></sub> | <i>A</i> <sup>2</sup> Δ <sub><i>u</i></sub> | <i>B</i> <sup>2</sup> Σ <sub><i>u</i></sub> <sup>+</sup> | <i>X</i> <sup>2</sup> Π       | <i>A</i> <sup>2</sup> Δ | <i>B</i> <sup>2</sup> Σ <sup>−</sup> | <i>C</i> <sup>2</sup> Σ <sup>+</sup> |
| SAC-CI-SDT- <i>R</i> /PS               | DZP   | 1.253                                       | 1.256                                       | 1.259                                                    | (1.400, 1.185)                | (1.315, 1.207)          | (1.302, 1.223)                       | (1.311, 1.214)                       |
| EA-EOMCCSD(2 <i>p</i> -1 <i>h</i> )    | DZP   | 1.259                                       | 1.258                                       | 1.260                                                    | (1.412, 1.196)                | (1.372, 1.186)          | (1.372, 1.190)                       | (1.365, 1.192)                       |
|                                        | x=D   | 1.260                                       | 1.258                                       | 1.260                                                    | (1.412, 1.193)                | (1.376, 1.182)          | (1.376, 1.186)                       | (1.375, 1.188)                       |
|                                        | x=T   | 1.242                                       | 1.245                                       | 1.247                                                    | (1.389, 1.178)                | (1.363, 1.166)          | (1.361, 1.170)                       | (1.356, 1.171)                       |
|                                        | x=Q   | 1.239                                       | 1.241                                       | 1.243                                                    | (1.385, 1.174)                | (1.362, 1.162)          | (1.360, 1.166)                       | (1.356, 1.167)                       |
| EA-EOMCCSD(3 <i>p</i> -2 <i>h</i> )    | DZP   | 1.261                                       | 1.262                                       | 1.264                                                    | (1.410, 1.198)                | (1.329, 1.217)          | (1.308, 1.241)                       | (1.322, 1.224)                       |
|                                        | x=D   | 1.262                                       | 1.261                                       | 1.263                                                    | (1.409, 1.195)                | (1.332, 1.212)          | (1.313, 1.234)                       | (1.325, 1.220)                       |
|                                        | x=T   | 1.246                                       | 1.246                                       | 1.248                                                    | (1.388, 1.180)                | (1.316, 1.196)          | (1.297, 1.215)                       | (1.308, 1.203)                       |
| EA-EOMCCSD(3 <i>p</i> -2 <i>h</i> ){4} | DZP   | 1.262                                       | 1.262                                       | 1.264                                                    | (1.411, 1.197)                | (1.329, 1.217)          | (1.308, 1.241)                       | (1.322, 1.224)                       |
|                                        | x=D   | 1.262                                       | 1.261                                       | 1.264                                                    | (1.408, 1.195)                | (1.332, 1.212)          | (1.313, 1.234)                       | (1.325, 1.220)                       |
|                                        | x=T   | 1.246                                       | 1.246                                       | 1.249                                                    | (1.389, 1.180)                | (1.316, 1.196)          | (1.297, 1.216)                       | (1.308, 1.203)                       |
|                                        | x=Q   | 1.242                                       | 1.243                                       | 1.245                                                    | (1.387, 1.175)                | (1.315, 1.190)          | (1.295, 1.210)                       | (1.307, 1.197)                       |
| Experiment <sup>c</sup>                |       | 1.245                                       | 1.249                                       | 1.259                                                    |                               |                         |                                      |                                      |

<sup>a</sup> The R<sub>C-N</sub> bond lengths in Å. The *D*<sub>2*h*</sub> symmetry was employed.

<sup>b</sup> The numbers in parentheses report the R<sub>C-C</sub> and R<sub>C-N</sub> bond lengths, respectively, in Å. The *C*<sub>2*v*</sub> symmetry was employed.

<sup>c</sup> Taken from [195–197].

Having demonstrated the significance of higher than  $2p-1h$  contributions for an accurate description of the excitation energies in the CNC and C<sub>2</sub>N molecules and having established the ability of the active-space EA-EOMCCSD( $3p-2h$ ) approach to capture the most significant  $3p-2h$  contributions with only a few active orbitals, independent of the basis set, we turn now to the effectiveness of the EA-EOMCC schemes in describing the equilibrium geometries of the ground and excited states of CNC and C<sub>2</sub>N. As seen in Table 3.3, the EA-EOMCCSD( $2p-1h$ ) level of theory gives the C–N bond lengths in CNC, designated as  $R_{\text{C-N}}$ , which deviate from the corresponding experimental values by 0.003-0.015 Å for the  $X\ ^2\Pi_g$  state, 0.004-0.009 Å for the  $A\ ^2\Delta_u$  state, and 0.001-0.016 Å for the  $B\ ^2\Sigma_u^+$  state, when the cc-pV $x$ Z basis sets with  $x = \text{D, T, and Q}$  are employed. The full EA-EOMCCSD( $3p-2h$ ) approach employing the cc-pVDZ and cc-pVTZ basis sets produces the  $R_{\text{C-N}}$  values that deviate from experiment by 0.001-0.017 Å, 0.003-0.012 Å, and 0.004-0.011 Å for the  $X\ ^2\Pi_g$ ,  $A\ ^2\Delta_u$ , and  $B\ ^2\Sigma_u^+$  states, respectively, i.e., the results that are of equally high quality and not much different than the low-order EA-EOMCCSD( $2p-1h$ ) data. The analogous active-space EA-EOMCCSD( $3p-2h$ ){4} calculations give the  $R_{\text{C-N}}$  bond lengths that differ from experiment by 0.001-0.017 Å in the case of the  $X\ ^2\Pi_g$  state, 0.003-0.012 Å in the case of the  $A\ ^2\Delta_u$  state, and 0.005-0.014 Å when the  $B\ ^2\Sigma_u^+$  state is examined. All of this shows that not only is the EA-EOMCCSD( $3p-2h$ ){4} approach able to reproduce the more computationally demanding EA-EOMCCSD( $3p-2h$ ) results for the nuclear geometries of the low-lying states of CNC, but that the differences between the high-level EA-EOMCCSD( $3p-2h$ ) values of  $R_{\text{C-N}}$  and those obtained with the inexpensive EA-EOMCCSD( $2p-1h$ ) method differ only by 0.002-0.004 Å for the  $X\ ^2\Pi_g$  state and 0.001-0.003 Å for the  $A\ ^2\Delta_u$  and

$B\ ^2\Sigma_u^+$  states, at least when the cc-pVDZ and cc-pVTZ basis sets are employed. The active-space EA-EOMCCSD( $3p-2h$ ){4} approach and the EA-EOMCCSD( $2p-1h$ ) method give the  $R_{C-N}$  values that differ by at most 0.004 Å for all states of CNC and all basis sets examined in this work, confirming the observation that it is sufficient to use the low-level EA-EOMCCSD( $2p-1h$ ) approach to obtain an accurate description of the equilibrium geometries of the low-lying states of CNC.

Similar, but not entirely identical, remarks apply to the  $C_2N$  molecule. As shown in Table 3.3, the EA-EOMCCSD( $2p-1h$ ) and EA-EOMCCSD( $3p-2h$ ){4} approaches employing the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets give the C–C and C–N bond lengths,  $R_{C-C}$  and  $R_{C-N}$ , respectively, that differ by at most 0.065 Å when we compare the EA-EOMCCSD( $2p-1h$ ) and the corresponding EA-EOMCCSD( $3p-2h$ ){4} data for all electronic states of  $C_2N$  examined in this work, mostly because of the inability of the EA-EOMCCSD( $2p-1h$ ) approach to provide a highly accurate description of the excited-state geometries [the differences between the EA-EOMCCSD( $2p-1h$ ) and EA-EOMCCSD( $3p-2h$ ){4} geometries of the  $C_2N$ ’s ground state are less than 0.004 Å]. On the other hand, the differences between the active-space EA-EOMCCSD( $3p-2h$ ){4} and full EA-EOMCCSD( $3p-2h$ ) values of  $R_{C-C}$  and  $R_{C-N}$  obtained with the cc-pVDZ and cc-pVTZ basis sets do not exceed 0.001 Å, confirming our earlier remarks about the ability of the active-space EA-EOMCCSD( $3p-2h$ ) approach to capture essentially all correlation effects that are included in the full EA-EOMCCSD( $3p-2h$ ) calculations. We could not find any experimental data for the geometries of the ground and excited states of  $C_2N$ , so we cannot comment on the accuracy of our  $R_{C-C}$  and  $R_{C-N}$  values resulting from the EA-EOMCC calculations in any

definitive manner, but, judging by the high quality of the EA-EOMCC results for the geometries of the low-lying states of CNC, we can conclude that the geometries resulting from the full EA-EOMCCSD( $3p-2h$ ) and active-space EA-EOMCCSD( $3p-2h$ ){4} calculations using the cc-pVTZ or cc-pVQZ basis sets are of the similarly high quality. The low-level EA-EOMCCSD( $2p-1h$ ) calculations seem less accurate than in the CNC case, particularly when the excited states of C<sub>2</sub>N are examined, but they are still in reasonable agreement with the high-level full and active-space EA-EOMCCSD( $3p-2h$ ) results.

The above discussion provides us with an important insight about the performance of the EA-EOMCC methods. The EA-EOMCCSD( $2p-1h$ ) approach, while generally inadequate for an accurate description of the excitation energies in open-shell systems, such as the CNC and C<sub>2</sub>N molecules examined in this work, is capable of providing reasonably accurate equilibrium geometries, even for excited states that have a significant multireference character. On the other hand, it seems to be generally safer to use the active-space EA-EOMCCSD( $3p-2h$ ) approach in geometry optimizations, particularly that it provides the results that are virtually identical to the corresponding full EA-EOMCCSD( $3p-2h$ ) data, both for the excitation energies and nuclear geometries.

### 3.2.2.2 IP-EOMCC results for NCO and N<sub>3</sub>

We now turn to the IP-EOMCC calculations for the NCO and N<sub>3</sub> molecules, which are summarized in Tables 3.4–3.6. For both molecules, the IP-EOMCC optimizations employing the DZP basis set, carried out in the present work, produced results that are very similar to the previously reported [52, 179] IP-EOMCC adiabatic excitation energies calculated at the SAC-CI-SDT-*R*/PS optimized geometries, all obtained with the same DZP basis as that



used here. For example, the deviations between the adiabatic excitation energies of NCO and N<sub>3</sub> obtained in the present IP-EOMCC/DZP optimizations and the analogous excitation energies reported in Refs. [52,179], which used the geometries obtained with the SAC-CI-SDT-*R*/PS approach, are 0.001-0.078 eV for all states and all methods considered in this study. As in the case of the EA-EOMCC calculations discussed in Section 3.2.2.1, after seeing that our present IP-EOMCC optimizations for the ground and excited states of NCO and N<sub>3</sub> employing the DZP basis set were able to reproduce the analogous results calculated at the geometries obtained in the SAC-CI-SDT-*R*/PS calculations, reported in Refs. [52,179], we moved to the examination of the effect of the use of the correlation-consistent basis sets of the cc-pV $x$ Z quality on the calculated excitation energies and geometries. As explained in Section 3.2.1, we exploited the  $C_{2v}$  symmetry in the calculations for NCO and the  $D_{2h}$  symmetry in the calculations for N<sub>3</sub>. The  $X\ ^2\Pi$  and  $B\ ^2\Pi$  states of NCO were calculated as the two lowest-energy  $^2B_1$  states, although we could also use the two lowest-energy  $^2B_2$  states, obtaining the same results. The  $A\ ^2\Sigma^+$  state of NCO was obtained as the lowest-energy  $^2A_1$  state. The  $X\ ^2\Pi_g$  state of N<sub>3</sub> was calculated as the lowest-energy  $^2B_{2g}$  state (we could use the lowest-energy  $^2B_{3g}$  state instead and obtain identical results) and the  $B\ ^2\Sigma_u^+$  of N<sub>3</sub> was calculated as the lowest-energy  $^2B_{1u}$  state. The discussion provided below focuses on the IP-EOMCC calculations using the cc-pV $x$ Z ( $x = D, T, Q$ ) basis sets.

Table 3.4: Total and adiabatic excitation energies for the ground and low-lying excited states of NCO, as obtained with the different IP-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ (x=D,T,Q) basis sets and extrapolating to the CBS limit.

| Method                   | Basis | Total Energy (hartree) |               |             | Adiabatic Excitation Energy (eV) |                 |
|--------------------------|-------|------------------------|---------------|-------------|----------------------------------|-----------------|
|                          |       | $X^2\Pi$               | $A^2\Sigma^+$ | $B^2\Pi$    | $A^2\Sigma^+-X^2\Pi$             | $B^2\Pi-X^2\Pi$ |
| IP-EOMCCSD( $2h-1p$ )    | DZP   | -167.581951            | -167.475380   | -167.427707 | 2.900                            | 4.197           |
|                          | x=D   | -167.576116            | -167.468912   | -167.419125 | 2.919                            | 4.273           |
|                          | x=T   | -167.718401            | -167.613444   | -167.560443 | 2.856                            | 4.298           |
|                          | x=Q   | -167.763112            | -167.659168   | -167.604432 | 2.828                            | 4.318           |
|                          | CBS-A | -167.786913            | -167.683604   | -167.626529 | 2.811                            | 4.364           |
|                          | CBS-B | -167.788412            | -167.685072   | -167.629275 | 2.812                            | 4.330           |
| IP-EOMCCSD( $3h-2p$ )    | DZP   | -167.591701            | -167.486508   | -167.448441 | 2.862                            | 3.898           |
|                          | x=D   | -167.587630            | -167.481331   | -167.441319 | 2.893                            | 3.981           |
|                          | x=T   | -167.732789            | -167.628442   | -167.584981 | 2.839                            | 4.022           |
| IP-EOMCCSD( $3h-2p$ ){2} | DZP   | -167.589579            | -167.476255   | -167.446490 | 3.081                            | 3.891           |
|                          | x=D   | -167.585489            | -167.470340   | -167.438481 | 3.133                            | 4.000           |
|                          | x=T   | -167.730109            | -167.617771   | -167.581463 | 3.057                            | 4.045           |
|                          | x=Q   | -167.775958            | -167.664698   | -167.626155 | 3.028                            | 4.076           |
|                          | CBS-A | -167.799482            | -167.687223   | -167.648439 | 3.055                            | 4.110           |
|                          | CBS-B | -167.801944            | -167.691316   | -167.651415 | 3.010                            | 4.096           |
| Experiment <sup>a</sup>  |       |                        |               |             | 2.821                            | 3.937           |

<sup>a</sup> Taken from [195].

Unlike the CNC and C<sub>2</sub>N molecules, which are characterized by the presence of the low-lying excited states with a significant multireference character in their respective electronic spectra, the low-lying states of NCO and N<sub>3</sub> have a predominantly  $1h$  excitation character relative to the corresponding NCO<sup>-</sup> and N<sub>3</sub><sup>-</sup> reference ions, with only small contributions from higher-than- $1h$  excitations. As a result, it is much easier to describe the low-lying states of NCO and N<sub>3</sub> by the IP-EOMCC methods and already the basic IP-EOMCCSD( $2h-1p$ ) approach performs quite well. For example, as shown in Table 3.4, the deviations from experiment for the adiabatic excitations in NCO resulting from the IP-EOMCCSD( $2h-1p$ ) calculations are only 0.007-0.098 eV for the  $A\ ^2\Sigma^+$  state and 0.336-0.381 eV for the  $B\ ^2\Pi$  state when the cc-pV $x$ Z basis sets with  $x = D, T,$  and  $Q$  are employed. Inclusion of higher-order ( $3h-2p$ ) correlation effects through the full IP-EOMCCSD( $3h-2p$ ) method offers additional improvements, reducing the overall deviations from experiment to 0.018-0.072 eV in the  $A\ ^2\Sigma^+$  case and 0.044-0.085 eV in the  $B\ ^2\Pi$  case, when the cc-pVDZ and cc-pVTZ basis sets are employed.

The inexpensive active-space variant of IP-EOMCCSD( $3h-2p$ ) using only two active occupied orbitals, IP-EOMCCSD( $3h-2p$ ){2}, yields similar excitation energy values as its more expensive parent scheme, with somewhat larger deviations from experiment of 0.207-0.312 eV for the  $A\ ^2\Sigma^+$  state and very small 0.063-0.139 eV deviations for the  $B\ ^2\Pi$  state, confirming that one can essentially use any IP-EOMCC approach and obtain a reasonable description of the low-lying states of NCO, but the deviations between the results of the full and active-space IP-EOMCCSD( $3h-2p$ ) calculations for NCO are somewhat larger than those observed in the EA-EOMCCSD( $3p-2h$ ) computations for CNC and C<sub>2</sub>N. This is particularly true

for the  $A\ ^2\Sigma^+$  state, where they are 0.240 eV for the adiabatic excitation energy and 11.0 millihartree for the total energy when the cc-pVDZ basis set is employed and 0.218 eV and 10.7 millihartree when the cc-pVTZ basis set is used. As pointed out in previous work [179], these larger differences between the full and active-space IP-EOMCCSD( $3h-2p$ ) results for the  $A\ ^2\Sigma^+$  state of NCO are likely due to the relatively small active space used in the latter calculations, which consists of only one pair of highest-energy occupied  $\pi$  orbitals of  $\text{NCO}^-$ , and/or from changes in the character of molecular orbitals when going from the  $\text{NCO}^-$  reference ion to the NCO target species. On the other hand, the overall agreement between the full and active-space IP-EOMCCSD( $3h-2p$ ) results for NCO is rather good. For example, the differences between the full and active-space IP-EOMCCSD( $3h-2p$ ) results for the adiabatic excitation energies corresponding to the  $B\ ^2\Pi$  state are only 0.019 eV when the cc-pVDZ basis set is employed and 0.023 eV when the cc-pVTZ basis set is used. The differences between the total energies obtained in the full IP-EOMCCSD( $3h-2p$ ) and active-space IP-EOMCCSD( $3h-2p$ ){2} calculations for the  $X\ ^2\Pi$  and  $B\ ^2\Pi$  states range between 2.1 and 3.5 millihartree when the cc-pVDZ and cc-pVTZ basis set are used, which is an excellent agreement.

Many of the above observations remain valid when the IP-EOMCC methods are applied to  $\text{N}_3$ . As shown in Table 3.5, the adiabatic excitation energies corresponding to the  $B\ ^2\Sigma_u^+$  state obtained with the IP-EOMCCSD( $2h-1p$ ) optimizations employing the cc-pV $x$ Z basis sets with  $x = \text{D, T, and Q}$  differ from the corresponding experimental value by 0.056-0.110 eV. Again, as in the NCO case, the full IP-EOMCCSD( $3h-2p$ ) approach reduces the already small errors in the IP-EOMCCSD( $2h-1p$ ) results for the  $B\ ^2\Sigma_u^+$  state of  $\text{N}_3$  to the even smaller

0.023-0.049 eV range. The much less expensive active-space IP-EOMCCSD( $3h-2p$ ){2} calculations using the cc-pV $x$ Z basis sets with  $x = \text{D, T, and Q}$  produce the 0.174-0.200 eV errors, which are larger than those obtained with full IP-EOMCCSD( $3h-2p$ ), but the general agreement between the full and active-space IP-EOMCCSD( $3h-2p$ ) data is reasonable. Indeed, the total energies resulting from the full IP-EOMCCSD( $3h-2p$ ) and active space IP-EOMCCSD( $3h-2p$ ){2} calculations differ by only 1.4-1.7 millihartree in the  $X\ ^2\Pi_g$  case and 6.9-7.3 millihartree in the case of the  $B\ ^2\Sigma_u^+$  state. The adiabatic excitation energies corresponding to the  $B\ ^2\Sigma_u^+$  state obtained in the full and active-space IP-EOMCCSD( $3h-2p$ ) calculations with the cc-pVDZ and cc-pVTZ basis sets differ by 0.151 eV, which is a reasonable agreement. Again, the somewhat larger differences between the full and active-space IP-EOMCCSD( $3h-2p$ ) data for the  $B\ ^2\Sigma_u^+$  state compared to the analogous EA-EOMCC calculations for CNC and C<sub>2</sub>N are likely due to the reasons cited above for the NCO molecule.

Table 3.5: Total and adiabatic excitation energies for the ground and low-lying excited states of N<sub>3</sub>, as obtained with the different IP-EOMCC approaches using the DZP [4s2p1d] and cc-pVxZ (x=D,T,Q) basis sets and extrapolating to the CBS limit.

| Method                   | Basis | Total Energy (hartree) |                   | Adiabatic Excitation Energy (eV) |
|--------------------------|-------|------------------------|-------------------|----------------------------------|
|                          |       | $X\ ^2\Pi_g$           | $B\ ^2\Sigma_u^+$ | $B\ ^2\Sigma_u^+ - X\ ^2\Pi_g$   |
| IP-EOMCCSD( $2h-1p$ )    | DZP   | -163.716374            | -163.545829       | 4.641                            |
|                          | x=D   | -163.712083            | -163.542627       | 4.611                            |
|                          | x=T   | -163.848768            | -163.677460       | 4.662                            |
|                          | x=Q   | -163.891293            | -163.719861       | 4.665                            |
|                          | CBS-A | -163.923747            | -163.752426       | 4.662                            |
|                          | CBS-B | -163.915306            | -163.743856       | 4.665                            |
| IP-EOMCCSD( $3h-2p$ )    | DZP   | -163.729782            | -163.560803       | 4.598                            |
|                          | x=D   | -163.726673            | -163.558437       | 4.578                            |
|                          | x=T   | -163.865416            | -163.696218       | 4.604                            |
| IP-EOMCCSD( $3h-2p$ ){2} | DZP   | -163.728362            | -163.554434       | 4.733                            |
|                          | x=D   | -163.725315            | -163.551533       | 4.729                            |
|                          | x=T   | -163.863766            | -163.689041       | 4.755                            |
|                          | x=Q   | -163.907325            | -163.732710       | 4.752                            |
|                          | CBS-A | -163.939566            | -163.764109       | 4.747                            |
|                          | CBS-B | -163.931977            | -163.757469       | 4.749                            |
| Experiment <sup>a</sup>  |       |                        |                   | 4.555                            |

<sup>a</sup> Taken from [195].

We now turn our attention to the numerical stability of our IP-EOMCC results for the NCO and N<sub>3</sub> molecules in the CBS limit. As in the EA-EOMCC calculations for CNC and C<sub>2</sub>N, the IP-EOMCC total energies of each state of NCO and N<sub>3</sub> shown in Tables 3.4 and 3.5 behave in a systematic manner, as we go from the cc-pVDZ to cc-pVQZ basis sets, showing the initial signs of convergence, and the excitation energies obtained in the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ){2} calculations with the cc-pVQZ basis set can be regarded as reasonably well converged, which helps the validity of the CBS-A extrapolations. Indeed, the differences between the adiabatic excitation energies calculated with the cc-pVTZ and cc-pVQZ basis sets at the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ){2} levels of theory are 0.028-0.029 eV for the  $A\ ^2\Sigma^+$  state of NCO, 0.020-0.031 eV for the  $B\ ^2\Pi$  state of NCO, and 0.003 eV for the  $B\ ^2\Sigma_u^+$  state of N<sub>3</sub>. It is, therefore, not surprising that the CBS-A and CBS-B extrapolations for the ground and excited states of the NCO and N<sub>3</sub> molecules summarized in Tables 3.4 and 3.5 are in good agreement. Indeed, the CBS-A and CBS-B total energies obtained with the IP-EOMCCSD( $2h-1p$ ) data for NCO differ by only 1.5 millihartree for the  $X\ ^2\Pi$  and  $A\ ^2\Sigma^+$  states and 2.7 millihartree for the  $B\ ^2\Pi$  state. The corresponding excitation energies resulting from both CBS extrapolations differ by 0.001 eV for the  $A\ ^2\Sigma^+$  state and 0.034 eV for the  $B\ ^2\Pi$  state. In consequence, the CBS-A- and CBS-B-extrapolated IP-EOMCCSD( $2h-1p$ ) excitation energies obtained for NCO differ from the corresponding experimental values by 0.009-0.010 eV in the  $A\ ^2\Sigma^+$  case and 0.393-0.427 eV in the case of the  $B\ ^2\Pi$  state. Similar remarks apply to the IP-EOMCCSD( $3h-2p$ ){2} approach, where the corresponding CBS-A- and CBS-B-extrapolated total energies differ by 2.5, 4.1, and 3.0 millihartree for the  $X\ ^2\Pi$ ,  $A\ ^2\Sigma^+$ , and  $B\ ^2\Pi$  states, respectively, so

that the differences in the resulting CBS-A and CBS-B IP-EOMCCSD( $3h-2p$ ){2} excitation energies are 0.045 eV for the  $A\ ^2\Sigma^+$  state and 0.014 eV for the  $B\ ^2\Pi$  state. As a consequence, the CBS-A- and CBS-B-extrapolated IP-EOMCCSD( $3h-2p$ ){2} adiabatic excitation energies for NCO differ from experiment by 0.189-0.234 eV for the  $A\ ^2\Sigma^+$  state and 0.159-0.173 eV for the  $B\ ^2\Pi$  state. Much of the above analysis applied to  $N_3$ . Indeed, although the CBS-A and CBS-B extrapolations applied to the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ){2} total energies produce somewhat larger differences than in the case of NCO (7.6-8.4 millihartree in the case of the  $X\ ^2\Pi_g$  state and 6.6-8.6 millihartree in the case of the  $B\ ^2\Sigma_u^+$  state), the adiabatic excitation energies resulting from both CBS extrapolations are very stable, to within 0.003 eV for the IP-EOMCCSD( $2h-1p$ ) approach and 0.002 eV for the IP-EOMCCSD( $3h-2p$ ){2} method. The CBS-A- and CBS-B-extrapolated IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ){2} adiabatic excitation energies corresponding to the  $B\ ^2\Sigma_u^+$  state of  $N_3$  are within 0.107-0.194 eV from experiment.

To conclude our discussion, we look at the performance of the IP-EOMCC methods in describing the equilibrium geometries of the ground and low-lying excited states of the NCO and  $N_3$  species examined in this work. The results of our geometry optimizations for NCO and  $N_3$  are summarized in Table 3.6. In the case of the  $X\ ^2\Pi$  state of the NCO molecule, the basic IP-EOMCCSD( $2h-1p$ ) approach produces results that deviate from experiment by 0.016-0.031 Å for the N–C bond length (designated as  $R_{N-C}$ ) and 0.018-0.033 Å for the C–O bond length (designated as  $R_{C-O}$ ) when the cc-pV $x$ Z basis sets with  $x = D, T,$  and  $Q$  are employed. The same approach applied to the  $A\ ^2\Sigma^+$  state of NCO gives the 0.014-0.031 Å errors for  $R_{N-C}$  and 0.018-0.031 Å for  $R_{C-O}$ . The IP-EOMCCSD( $3h-2p$ )



results exhibit very similar trends and accuracies, confirming the small role of higher-order contributions neglected in IP-EOMCCSD( $2h-1p$ ) and present in IP-EOMCCSD( $3h-2p$ ). The differences between the IP-EOMCCSD( $3h-2p$ ) and experimental values of  $R_{\text{N-C}}$  are 0.024-0.038 Å for the  $X\ ^2\Pi$  state and 0.021-0.035 Å for the  $A\ ^2\Sigma^+$  state. The analogous differences for  $R_{\text{C-O}}$  are 0.015-0.025 Å for the  $X\ ^2\Pi$  state and 0.012-0.022 Å for the  $A\ ^2\Sigma^+$  state. Although it may very well be that higher-than- $3h-2p$  contributions neglected in the IP-EOMCCSD( $3h-2p$ ) calculations and high angular momentum functions that are not present in the cc-pVTZ (or cc-pVQZ) basis sets are the sources of the above errors, it is also possible that the experimental geometries of the  $X\ ^2\Pi$  and  $A\ ^2\Sigma^+$  states of NCO reported in Ref. [195] might be in some error too, since none of the states of NCO examined here is as challenging as some of the states of CNC and C<sub>2</sub>N discussed in Section 3.2.2.1. While there are unexplained differences between the experimentally determined N-C and C-O bond lengths in the  $X\ ^2\Pi$  and  $A\ ^2\Sigma^+$  states of NCO and our theoretical predictions, it is of great interest to note that the differences between the results of the geometry optimizations using the full and active-space IP-EOMCCSD( $3h-2p$ ) approaches are virtually none. Indeed, there is no difference (to within 0.001 Å) between the full IP-EOMCCSD( $3h-2p$ ) and active-space IP-EOMCCSD( $3h-2p$ ){2} results for the N-C bond length in the  $X\ ^2\Pi$  state and the corresponding C-O bond lengths differ by 0.002 Å only, when the cc-pVDZ and cc-pVTZ are employed. In the case of the  $A\ ^2\Sigma^+$  state, the differences between the full IP-EOMCCSD( $3h-2p$ ) and active-space IP-EOMCCSD( $3h-2p$ ){2} values of  $R_{\text{N-C}}$  and  $R_{\text{C-O}}$  are 0.004 Å and 0.001-0.002 Å, respectively. In the case of the  $B\ ^2\Pi$  state, these differences are 0.003 Å for  $R_{\text{N-C}}$  and 0.006-0.008 Å for  $R_{\text{C-O}}$ . The active-space IP-EOMCCSD( $3h-2p$ ){2}

calculations are clearly capable of reproducing the parent IP-EOMCCSD( $3h-2p$ ) data for the N–C and C–O bond lengths in the ground and excited states of NCO to very high accuracy.

Table 3.6: Comparison of the optimized equilibrium geometries for the low-lying states of  $\text{N}_3$  and  $\text{NCO}$ , as obtained with the IP-EOMCC and SAC-CI-SDT- $R$ /PS approaches using the DZP [4s2p1d] and cc-pVxZ (x=D,T,Q) basis sets.

| Method                   | Basis | $\text{N}_3^a$ |                    | $\text{NCO}^b$ |                  |                |
|--------------------------|-------|----------------|--------------------|----------------|------------------|----------------|
|                          |       | $X \ ^2\Pi_g$  | $B \ ^2\Sigma_u^+$ | $X \ ^2\Pi$    | $A \ ^2\Sigma^+$ | $B \ ^2\Pi$    |
| SAC-CI-SDT- $R$ /PS      | DZP   | 1.188          | 1.185              | (1.230, 1.193) | (1.191, 1.190)   | (1.220, 1.309) |
| IP-EOMCCSD( $2p-1h$ )    | DZP   | 1.195          | 1.191              | (1.232, 1.196) | (1.197, 1.192)   | (1.225, 1.318) |
|                          | x=D   | 1.185          | 1.181              | (1.231, 1.188) | (1.196, 1.184)   | (1.223, 1.313) |
|                          | x=T   | 1.171          | 1.169              | (1.219, 1.177) | (1.182, 1.175)   | (1.206, 1.306) |
|                          | x=Q   | 1.168          | 1.165              | (1.216, 1.173) | (1.179, 1.171)   | (1.202, 1.304) |
| IP-EOMCCSD( $3p-2h$ )    | DZP   | 1.200          | 1.196              | (1.240, 1.198) | (1.200, 1.198)   | (1.235, 1.328) |
|                          | x=D   | 1.190          | 1.187              | (1.238, 1.191) | (1.200, 1.190)   | (1.233, 1.322) |
|                          | x=T   | 1.175          | 1.173              | (1.224, 1.181) | (1.186, 1.180)   | (1.216, 1.312) |
| IP-EOMCCSD( $3p-2h$ ){2} | DZP   | 1.200          | 1.194              | (1.240, 1.196) | (1.196, 1.196)   | (1.239, 1.319) |
|                          | x=D   | 1.190          | 1.184              | (1.238, 1.189) | (1.196, 1.189)   | (1.236, 1.314) |
|                          | x=T   | 1.175          | 1.171              | (1.224, 1.179) | (1.182, 1.178)   | (1.219, 1.306) |
|                          | x=Q   | 1.172          | 1.168              | (1.222, 1.174) | (1.179, 1.174)   | (1.214, 1.303) |
| Experiment <sup>c</sup>  |       | 1.188          | 1.180              | (1.200, 1.206) | (1.165, 1.202)   |                |

<sup>a</sup> The  $R_{\text{N-N}}$  bond lengths in Å. The  $D_{2h}$  symmetry was employed.

<sup>b</sup> The numbers in parentheses report the  $R_{\text{N-C}}$  and  $R_{\text{C-O}}$  bond lengths, respectively, in Å. The  $C_{2v}$  symmetry was employed.

<sup>c</sup> Taken from [195].

Much of the above discussion applies to  $\text{N}_3$ . The nearest-neighbor N–N bond lengths, designated as  $R_{\text{N-N}}$ , resulting from the IP-EOMCCSD( $2h-1p$ ) calculations with the cc-pV $x$ Z basis sets with  $x = \text{D, T, and Q}$ , differ from the corresponding experimental data by 0.003-0.020 Å for the  $X\ ^2\Pi_g$  state and 0.001-0.015 Å for the  $B\ ^2\Sigma_u^+$  state. The higher-order IP-EOMCCSD( $3h-2p$ ) optimizations with the cc-pVDZ and cc-pVTZ basis set produce similar results, the 0.002-0.013 Å errors for the  $X\ ^2\Pi_g$  state and 0.007 Å for the  $B\ ^2\Sigma_u^+$  state. The active-space IP-EOMCCSD( $3h-2p$ ){2} approach, for which we could also afford the calculations with the cc-pVQZ basis set, produces the  $R_{\text{N-N}}$  values that deviate from experiment by 0.002-0.016 Å for the  $X\ ^2\Pi_g$  state and 0.004-0.012 Å for the  $B\ ^2\Sigma_u^+$  state. Again, there is a virtually perfect agreement between the expensive full IP-EOMCCSD( $3h-2p$ ) and relatively inexpensive active-space IP-EOMCCSD( $3h-2p$ ){2} calculations, where there is no difference (to within 0.001 Å) between the two sets of data in the case of the  $X\ ^2\Pi_g$  state and a very small, 0.002-0.003 Å, difference between the full IP-EOMCCSD( $3h-2p$ ) and active-space IP-EOMCCSD( $3h-2p$ ){2} values of  $R_{\text{N-N}}$  in the case of the  $B\ ^2\Sigma_u^+$  state. As in the case of the NCO molecule, the origin of the deviations between the IP-EOMCC calculations employing basis sets as large as cc-pVQZ, which seem numerically quite stable, and experimental  $R_{\text{N-N}}$  values could lie in the higher-than- $3h-2p$  correlations that we do not consider in this work or in the significance of the high angular momentum functions absent in the cc-pVTZ and cc-pVQZ bases, but one cannot exclude the possibility that the experimental data reported in Ref. [195] may need to be revisited. As in the EA-EOMCC calculations for the CNC and  $\text{C}_2\text{N}$  discussed in Section 3.2.2.1, it seems to us that the basic IP-EOMCCSD( $2h-1p$ ) method is capable of producing the optimized geometries of the ground- and excited-state NCO and

$\text{N}_3$  molecules that are comparable to those obtained with the computationally more demanding IP-EOMCCSD( $3h-2p$ ) methods, which is a useful observation from the point of view of other applications of such methods to geometry optimizations in other open-shell species.

### 3.3 Coupled-cluster interpretation of the photoelectron spectrum of $\text{Au}_3^-$

#### 3.3.1 Introductory remarks

In the previous section we showed that the EA- and IP-EOMCC methodologies can accurately describe the electronic states of challenging open-shell molecules consisting of second-row atoms. There also are many challenging open-shell transition metal species, whose various properties, including optical and photoelectron spectra, still await a satisfactory experimental and/or theoretical explanation. As part of this thesis, we undertook the study of the photoelectron spectrum of the small  $\text{Au}_3^-$  nanocluster, providing a comprehensive interpretation of this small, but very challenging molecule.

Since the pioneering works of Hutchings [199] and Haruta and coworkers [200] demonstrating the catalytic behavior of nanometer-sized gold particles, interest in and the study of their structural and electronic properties has steadily increased by both experimentalists and theorists [201–207]. While density functional theory (DFT) has been the workhorse for the majority of computations in this area (cf., e.g., Refs. [204–207]), the analogous investigations using high-level *ab initio* wave function approaches, particularly those based on CC theory remain rare due, in part, to prohibitive computational costs involved. There have

been a few investigations using the CCSD(T) approach, which examined the ground-state properties of the cation [208, 209], neutral [208, 210–212], and anion [208, 209, 211]  $\text{Au}_3$  isomers, with the latter two species being relevant to this work. Varganov *et al.* studied the absorption and activation of  $\text{O}_2$  [211] and  $\text{H}_2$  [213] on small  $\text{Au}_n$  and  $\text{Au}_n^-$  ( $n = 2, 3$ ) clusters using DFT, second-order Møller-Plesset perturbation theory (MP2), and CCSD(T). This has been followed by a number of CCSD(T) calculations for the low-lying isomers of  $\text{Au}_6$ ,  $\text{Au}_8$ , and  $\text{Au}_{10}$ , as well as  $\text{Au}_7^-$  comparing the results with MP2 and DFT [183, 209, 214–218]. Most recently, as discussed in Section 4.3, we studied the aerobic oxidation of methanol to formic acid on  $\text{Au}_8^-$ , comparing DFT computations with the CR-CC(2,3) approach [181]. The excitation spectra for the closed-shell  $\text{Au}_n$  clusters ( $n = 4, 6, 8$ ) were examined [219] using the EOMCCSD approach, obtaining an accurate description of one-electron transitions compared to experiment [220]. Much less is known, however, about the performance of the EOMCC theory when describing excited states of open-shell gold clusters.

As part of this dissertation, we undertook a thorough *ab initio* study of the photoelectron spectrum of the gold trimer anion [221–223],  $\text{Au}_3^-$ , by examining the ground and excited states of the corresponding neutral particle,  $\text{Au}_3$ , employing the scalar relativistic IP-EOMCC approximations (for the representative earlier *ab initio* calculations for  $\text{Au}_3$  and  $\text{Au}_3^-$ , see Refs. [208–212, 224–228]). We investigated [184] the effects of the basis set size, number of correlated electrons, level of applied theory, and geometry relaxation. As explained in Section 3.1, the IP-EOMCC family of methods used in our examination of the photoelectron spectrum of  $\text{Au}_3^-$  allow one to properly account for spin symmetry of open-shell species that the conventional CC/EOMCC approaches have difficulties with,

determining the electronic spectrum of the  $(N - 1)$ -electron system (in our case,  $\text{Au}_3$ ) by applying the linear ionizing operator,  $R_\mu^{(N-1)}$ , to the ground state of the corresponding  $N$ -electron closed-shell core (in our case,  $\text{Au}_3^-$ ) obtained with a single-reference CC approach. Our IP-EOMCC calculations for  $\text{Au}_3$  utilized two truncation schemes, namely, the basic IP-EOMCCSD( $2h-1p$ ) approximation and its higher-order IP-EOMCCSD( $3h-2p$ ) counterpart. Since the IP-EOMCCSD( $3h-2p$ ) calculations for the  $\text{Au}_3$  system using larger basis sets can be prohibitively expensive, especially if one decides to correlate semi-core electrons, in addition to the valence ones (as is the case in our study), we adopted a simple IP-EOMCC-based extrapolation scheme that captures the important higher-order  $3h-2p$  electron correlation contributions to the energy in an approximate, but computationally efficient manner. This extrapolation scheme is discussed first.

### 3.3.2 IP-EOMCC extrapolation scheme

To evaluate the effects of various factors entering IP-EOMCC computations, including the role of  $3h-2p$  excitations, basis set, and the number of correlated electrons, and determine their importance for providing accurate excitation and ionization energies of small open-shell gold clusters, such as  $\text{Au}_3$  and  $\text{Au}_3^-$ , we performed extensive calculations on  $\text{Au}^-$  and  $\text{Au}$ , varying the basis set, number of correlated electrons, and level of IP-EOMCC theory used to compute the ionization energies (IEs). The singlet ground state of the  $\text{Au}^-$  anion, which was used as the underlying closed-shell system out of which the open-shell ground and low-lying excited states of the  $\text{Au}$  atom were generated, was treated at the CCSD level of theory, with the corresponding orbitals of  $\text{Au}^-$  used to describe the electronic states of  $\text{Au}$ . We

first investigated the role of basis set size and level of IP-EOMCC theory in determining the ground and excited states of Au using the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ) approaches in conjunction with the aug-cc-pV $x$ Z-PP ( $x = D, T, Q, 5$ ) bases optimized for gold [229], correlating the  $5d^{10}6s^1$  valence electrons and an extra electron due to the charge in the CCSD calculation for  $\text{Au}^-$ . We then explored the effect of correlating the  $5s^25p^6$  semi-core or core-valence (CV) electrons by carrying out IP-EOMCCSD( $2h-1p$ )/aug-cc-pV $x$ Z-PP+CV and IP-EOMCCSD( $3h-2p$ )/aug-cc-pV $x$ Z-PP+CV ( $x = D, T, Q, 5$ ) calculations for the  $\text{Au}^-$  ion. We employed the scalar relativistic effective-core potential (ECP) created for use with the correlation-consistent basis sets, as originally developed in Ref. [230] and later modified to include the effects of the  $h$  functions in the pseudopotential in Ref. [231] to represent the effects of the core electrons of Au. For computations using the aug-cc-pVQZ-PP and aug-cc-pV5Z-PP basis sets the  $h$  and  $i$  functions were removed as GAMESS was not able to compute integrals involving functions with  $\ell \geq 5$  at the time. We extrapolated these results to the CBS limit again using the previously exploited formula given by Eq. (3.19), where this time  $x = 3, 4, 5$  in Eq. (3.19) is the cardinal number corresponding to the aug-cc-pVTZ-PP, aug-cc-pVQZ-PP, and aug-cc-pV5Z-PP basis sets, respectively,  $E(x)$  is the total CCSD energy of  $\text{Au}^-$  or the IP-EOMCC energy of the Au state of interest, computed with the aug-cc-pV $x$ Z-PP basis, and  $E_\infty$  is the total energy of the electronic state of  $\text{Au}^-$  or Au in the CBS limit. The results of these computations are presented in Table 3.7.



Table 3.7: Scalar relativistic ionization energies (in eV) of Au from Au<sup>-</sup> computed using the IP-EOMCC approaches; comparison with experiment and experimentally derived estimates.

| $2S+1L$ | $x$              | $2h-1p$ |          | $3h-2p$ |          | $1^a$ | $2^b$ | $2S+1L_J$   | B.E. <sup>c</sup> | Exp. Est. <sup>d</sup> | Scalar | Exp. Est. <sup>e</sup> |
|---------|------------------|---------|----------|---------|----------|-------|-------|-------------|-------------------|------------------------|--------|------------------------|
|         |                  | 5d6s    | 5s5p5d6s | 5d6s    | 5s5p5d6s |       |       |             |                   |                        |        |                        |
| $^2S$   | D                | 2.147   | 2.183    | 1.956   | 2.025    | 1.993 | 1.993 |             |                   |                        |        |                        |
|         | T                | 2.204   | 2.244    | 1.980   | 2.032    | 2.020 | 2.054 |             | 2.20 <sup>f</sup> |                        |        |                        |
|         | Q <sup>g</sup>   | 2.222   | 2.264    | 1.992   | 2.039    | 2.034 | 2.073 | $^2S_{1/2}$ |                   | 2.309                  |        | 2.309                  |
|         | 5 <sup>g</sup>   | 2.228   | 2.272    | 1.996   | 2.042    | 2.040 | 2.082 |             | 2.35              |                        |        |                        |
|         | CBS <sup>h</sup> | 2.231   | 2.277    | 1.998   | 2.044    | 2.043 | 2.086 |             |                   |                        |        |                        |
| $^2D$   | D                | 4.271   | 3.896    | 3.833   | 3.528    | 3.458 | 3.458 |             |                   |                        |        |                        |
|         | T                | 4.397   | 4.079    | 3.848   | 3.546    | 3.529 | 3.641 | $^2D_{5/2}$ | 3.45              | 3.444                  |        |                        |
|         | Q <sup>g</sup>   | 4.428   | 4.128    | 3.856   | 3.540    | 3.556 | 3.690 |             |                   |                        |        | 4.053                  |
|         | 5 <sup>g</sup>   | 4.438   | 4.144    | 3.859   | 3.537    | 3.565 | 3.706 | $^2D_{3/2}$ | 4.95              | 4.966                  |        |                        |
|         | CBS <sup>h</sup> | 4.444   | 4.153    | 3.860   | 3.535    | 3.570 | 3.715 |             |                   |                        |        |                        |
| $^2P^i$ | D                | 7.986   | 8.281    | 6.951   | 6.999    | 7.246 | 7.246 |             |                   |                        |        |                        |
|         | T                | 8.140   | 8.461    | 6.976   | 7.039    | 7.297 | 7.426 | $^2P_{1/2}$ |                   | 6.941                  |        |                        |
|         | Q <sup>g</sup>   | 8.167   | 8.493    | 6.993   | 7.058    | 7.319 | 7.458 |             |                   |                        |        | 7.256                  |
|         | 5 <sup>g</sup>   | 8.174   | 8.504    | 6.999   | 7.069    | 7.329 | 7.469 | $^2P_{3/2}$ |                   | 7.414                  |        |                        |
|         | CBS <sup>h</sup> | 8.177   | 8.510    | 7.002   | 7.075    | 7.335 | 7.475 |             |                   |                        |        |                        |

<sup>a</sup> Extrapolated values obtained when BS1 = BS2, using Eq. (3.19).

<sup>b</sup> Extrapolated values obtained when BS1 = aug-cc-pVDZ-PP and BS2 = aug-cc-pVxZ-PP, in Eq. (3.19).

<sup>c</sup> Experimental binding energies taken from Ref. [223].

<sup>d</sup> Experimental estimates calculated as the sum of the electron affinity [232] and the excitation energies of the states [233].

<sup>e</sup> Estimated scalar relativistic peak values determined from the theoretically derived experimental values using the formula

$\sum_J (2J+1) *^{2S+1}L_J / \sum_J (2J+1)$ , where  $^{2S+1}L_J$  is the energy of the relativistic term in a J coupled scheme.

<sup>f</sup> There is a shoulder in the 2.35 eV peak at 2.20 eV, which could also be contamination from other species, such as the (Au—Au)<sup>2-</sup> dimer.

<sup>g</sup> Basis functions with  $\ell \geq 5$  were removed from the basis set due to restrictions in the GAMESS program.

<sup>h</sup> Complete basis set (CBS) limit values determined using Eq. (3.19) with  $x = 3, 4, 5$ .

<sup>i</sup> This state lies outside the experimental range of 6.424 eV in Ref. [223].

While the augmented correlation-consistent basis sets are somewhat larger than their cc-pVxZ-PP counterparts and therefore increase the costs of the CC/EOMCC computations, we found that it was necessary to use the aug-cc-pVxZ-PP bases, particularly in the  $x = \text{D}$  and  $\text{T}$  cases, to obtain reasonable answers. As shown in Table 3.7, the aug-cc-pVDZ-PP basis is not capable of providing results of the quality needed to accurately describe the photoelectron spectra of Au containing species, with errors as large as  $\sim 0.3$  eV compared to the corresponding CBS data. When the  $x = \text{T}$  basis set is employed, there is a dramatic improvement in the results, with errors relative to the CBS values dropping to less than 0.08 eV, with the larger ( $x = \text{Q}, 5$ ) aug-cc-pVxZ-PP bases reducing the errors even further, though their improvements relative to the aug-cc-pVTZ-PP calculations are not nearly as large. Thus, by incorporating basis sets of the aug-cc-pVTZ-PP quality in IP-EOMCC calculations, we can safely assume the data are reasonably converged in terms of the basis set size.

While the IP-EOMCCSD( $2h-1p$ ) approach is capable of providing an accurate description of states dominated by  $1h$  contributions, it cannot accurately account for states with significant higher-order correlations ( $2h-1p$ ,  $3h-2p$ , etc.), which are quite common in transition metal clusters. This point is amply demonstrated by examining the IEs corresponding to the  $^2\text{P}$  state of Au in Table 3.7. Comparing the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ) results for the  $^2\text{P}$  state of Au, which has substantial  $2h-1p$  contributions to its wave function, we can see that the effect of  $3h-2p$  components of  $R_\mu^{(N-1)}$  is on the order of 1.5 eV, independent of the basis set employed. For the  $^2\text{S}$  and  $^2\text{D}$  states the effects of  $3h-2p$  correlations are smaller (about 0.2 eV for the  $^2\text{S}$  state and about 0.6 eV for

the  $^2D$  state), but they are still considerably larger than the changes due to the basis set size. Interestingly, however, by examining the IP-EOMCCSD( $3h-2p$ )/aug-cc-pVDZ-PP and IP-EOMCCSD( $3h-2p$ )/aug-cc-pVTZ-PP results, we can see that the contributions due to higher-order  $3h-2p$  correlations are accurately characterized with the smaller aug-cc-pVDZ-PP basis set. All of this is telling us that we need the aug-cc-pVTZ-PP basis set to capture the basic correlation effects included at the IP-EOMCCSD( $2h-1p$ ) level, but it is at the same time sufficient to employ the smaller aug-cc-pVDZ-PP basis to determine the effect of  $3h-2p$  correlations.

The role of correlating the CV electrons can also have a non-negligible effect on the IEs of Au. For the  $^2S$  state of Au, the calculated IEs change by less than 0.07 eV when CV correlations are included. On the other hand, for the  $^2D$  and  $^2P$  states they change by as much as 0.38 eV. Thus, while both the IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ) IEs are affected by inclusion of CV correlations, the magnitude of the change in IEs is relatively stable, and equal for both approaches, once basis sets of at least the aug-cc-pVTZ-PP quality are employed in the IP-EOMCC calculations. This will allow us to estimate the contributions of the CV effects for larger systems, such as Au<sub>3</sub>, using the lower-level IP-EOMCCSD( $2h-1p$ ) approach.

Taking the above discussion into consideration we conclude that to obtain an accurate description of IEs or excitation energies in gold clusters and to capture the most important correlation and basis set effects, it is sufficient to perform the lower-level IP-EOMCCSD( $2h-1p$ ) calculations using the aug-cc-pVTZ-PP basis set, correlating the valence and semi-core electrons, while estimating the role of  $3h-2p$  correlations using the IP-EOMCCSD( $3h-2p$ ) calcula-

tions with a smaller aug-cc-pVDZ-PP basis, which correlate valence electrons only. Thus, we can propose the following composite approach to provide an accurate and computationally efficient IP-EOMCC scheme for examining the photoelectron spectra of gold nanoparticles:

$$IE_k = [IE_k(3h-2p) - IE_k(2h-1p)]/BS1 + IE_k(2h-1p)/[BS2 + CV], \quad (3.20)$$

where the first term on the right-hand side,  $[IE_k(3h-2p) - IE_k(2h-1p)]/BS1$ , is the  $3h-2p$  contribution for the IE of the  $k$ -th electronic state, of the  $(N-1)$ -electron cluster, computed as the difference between the IP-EOMCCSD( $3h-2p$ ) and IP-EOMCCSD( $2h-1p$ ) IEs using the smaller BS1 basis set of the aug-cc-pVDZ-PP or higher quality, which is subsequently added to the IP-EOMCCSD( $2h-1p$ )/[BS2+CV] IE value correlating valence and semi-core electrons and using a basis set BS2 of equal or greater size than BS1 if BS1 equals aug-cc-pV $x$ Z-PP with  $x \geq 3$  or larger than BS1 if BS1 = aug-cc-pVDZ-PP.

To test the above composite approach, we first set BS2 = BS1 for BS1 = aug-cc-pV $x$ Z-PP ( $x = D, T, Q, 5$ ), labeled as extrapolation scheme 1 in Table 3.7, and compare our results to the true IP-EOMCCSD( $3h-2p$ )/[BS2+CV] results. As shown in Table 3.7 we can see that the extrapolated IEs agree reasonably well with the true IP-EOMCCSD( $3h-2p$ )/[BS2+CV] computations, differing by 0.2–0.3 eV at most, though typically the extrapolated results fare even more favorably when compared to their true values.

We then tested how well the above extrapolation scheme performs when BS1 = aug-cc-pVDZ-PP and BS2 = aug-cc-pV $x$ Z-PP ( $x = D, T, Q, 5$ ), which we call in Table 3.7 the extrapolation scheme 2. For the  $^2S$  state of Au, which is largely a  $1h$  state relative to  $Au^-$ , the extrapolation scheme 2 reproduces the results of the full IP-EOMCCSD( $3h-2p$ )/[BS2+CV]

calculations to within 0.04 eV. For the challenging  $^2\text{D}$  and  $^2\text{P}$  states, which have important CV contributions and more significant  $2h\text{-}1p$  character, the agreement is not as good, but we are still able to reproduce the results of the true IP-EOMCCSD( $3h\text{-}2p$ )/[BS2+CV] calculation to within 0.2–0.4 eV in a computationally feasible manner, obtaining a reasonably accurate description. Most notably, much of the accuracy of the IP-EOMCCSD( $3h\text{-}2p$ )/aug-cc-pV $x$ Z-PP+CV approach with  $x \geq 3$  can be captured when BS1 and BS2 in Eq. (3.20) are aug-cc-pVDZ-PP and aug-cc-pVTZ-PP, respectively. We can, thus, expect that a composite IP-EOMCC scheme based on Eq. (3.20), where BS1 = aug-cc-pVDZ-PP and BS2 = aug-cc-pVTZ-PP, should provide an accurate representation of high-level IP-EOMCCSD( $3h\text{-}2p$ ) results obtained with larger basis sets of aug-cc-pVTZ-PP or higher quality, correlating valence as well as semi-core electrons.

While we were not able to account for higher-order relativistic effects, such as spin-orbit coupling, directly, i.e., using IP-EOMCC, the scalar relativistic effects were accurately accounted for in the carefully optimized ECP, which was derived from numerical all-electron multiconfiguration Dirac-Hartree-Fock [234] calculations [230,231]. Thus, while we were not able to account for the splitting of degenerate energy levels (e.g.,  $^2\text{D}$  into the  $^2\text{D}_{5/2}$  and  $^2\text{D}_{3/2}$ ), the estimated experimental scalar relativistic results obtained through suitable term averaging (see footnote *e* in Table 3.7) are in good agreement with our calculated results, further suggesting that our composite IP-EOMCC method should work well for the larger  $\text{Au}_n^-$  clusters with odd  $n$  values, such as  $\text{Au}_3^-$ . Our IP-EOMCC results for the photoelectron spectrum of  $\text{Au}_3^-$  are discussed next.

### 3.3.3 Results

We started [184] by optimizing the geometry of the singlet ground state of  $\text{Au}_3^-$  using the coarse-grain finite-difference model available in the CIOpt program suite [235], described in further details in Section 5.1, at the scalar relativistic CCSD level employing the aug-cc-pVDZ-PP basis set, correlating the  $5d^{10}6s^1$  valence electrons of each gold atom and an extra electron due to the charge. As in the case of the  $\text{Au}^-/\text{Au}$  system examined in the previous section, the aug-cc-pV $x$ Z-PP ( $x = \text{D}, \text{T}$ ) bases employed in our study were combined with the scalar relativistic ECP designed for use with the correlation-consistent basis sets, taken from Refs. [230,231]. Our CCSD-optimized structure of  $\text{Au}_3^-$  has a linear,  $\text{D}_{\infty h}$  symmetry, where the distance between nearest-neighbor gold atoms,  $r_{\text{Au-Au}}$ , is 2.593 Å, agreeing well with previous CC [208,209], other *ab initio* [208,209,211,224–226], and DFT [209,211,223] geometries. Using this  $\text{Au}_3^-$  equilibrium geometry, we computed the ground and several excited states of  $\text{Au}_3$ . Following the lessons learned from the IP-EOMCC calculations for the Au atom discussed in Section 3.3.2, we applied the scalar relativistic IP-EOMCCSD( $2h-1p$ ) approach, exploiting the aug-cc-pV $x$ Z-PP ( $x = \text{D}, \text{T}$ ) bases and the accompanying ECP and to investigate the role of the CV correlation effects, the more extensive IP-EOMCCSD( $2h-1p$ )/aug-cc-pV $x$ Z-PP+CV calculations correlating the  $5s^25p^6$  semi-core and  $5d^{10}6s^1$  valence electrons of each gold atom plus the extra electron due to the charge. As already pointed out, while the IP-EOMCCSD( $2h-1p$ ) method can reasonably describe the  $(N - 1)$ -electron states dominated by  $1h$  excitations relative to the  $N$ -electron closed-shell core, it cannot accurately account for states with significant higher-order correlations ( $2h-1p$ ,  $3h-2p$ , etc.). Owing to the fact that some low-lying states of  $\text{Au}_3$  are characterized by significant higher-than- $1h$

components, we also performed higher-order IP-EOMCCSD( $3h$ - $2p$ ) calculations correlating the  $5d^{10}6s^1$  valence electrons and employing the smaller aug-cc-pVDZ-PP basis set, since, as shown in Section 3.3.2, it is sufficient to use the aug-cc-pVDZ-PP basis to estimate the effects of  $3h$ - $2p$  correlations. The resulting vertical IEs are shown in Table 3.8 and Fig. 3.2(a)-(c).

Table 3.8: Vertical<sup>a</sup> (V) and adiabatic<sup>b</sup> (Ad) IEs (in eV) of Au<sub>3</sub><sup>-</sup> with respect to its ground state computed using the scalar relativistic IP-EOMCCSD(2*h*-1*p*)/aug-cc-pV*x*Z-PP(+CV) (*x* = D, T) and IP-EOMCCSD(3*h*-2*p*)/aug-cc-pVDZ-PP approaches.

| Ionization |                                                                           | Geometry                      |           | 2 <i>h</i> -1 <i>p</i> |                       | 2 <i>h</i> -1 <i>p</i> |                       | 3 <i>h</i> -2 <i>p</i> | Extrapolated <sup>e</sup> | Experiment <sup>f</sup> |
|------------|---------------------------------------------------------------------------|-------------------------------|-----------|------------------------|-----------------------|------------------------|-----------------------|------------------------|---------------------------|-------------------------|
| Energy     | State of Au <sub>3</sub>                                                  | <i>r</i> <sub>Au-Au</sub> (Å) | θ(degree) | 5d6s <sup>c</sup>      | 5s5p5d6s <sup>d</sup> | 5d6s <sup>c</sup>      | 5s5p5d6s <sup>d</sup> | 5d6s <sup>c</sup>      |                           |                         |
|            |                                                                           |                               |           | <i>x</i> = D           |                       | <i>x</i> = T           |                       | <i>x</i> = D           |                           |                         |
| V          | <sup>2</sup> Σ <sub><i>u</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.593                         | 180       | 3.621                  | 3.683                 | 3.670                  | 3.735                 | 3.427                  | 3.541                     | A:3.89(0.08)            |
| Ad         | X <sub>0</sub> <sup>2</sup> B <sub>1</sub> (C <sub>2<i>v</i></sub> )      | 2.725                         | 67        | 3.539                  | 3.578                 | 3.571                  | 3.604                 | 3.341                  | 3.406                     |                         |
| Ad         | X <sub>1</sub> <sup>2</sup> B <sub>1</sub> (C <sub>2<i>v</i></sub> )      | 2.575                         | 142       | 3.562                  | 3.598                 | 3.614                  | 3.645                 | 3.371                  | 3.454                     |                         |
| V          | <sup>2</sup> Σ <sub><i>g</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.593                         | 180       | 4.781                  | 4.546                 | 4.857                  | 4.709                 | 4.520                  | 4.448                     | B:4.38(0.15)            |
| Ad         | <sup>2</sup> Σ <sub><i>g</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.500                         | 180       | 4.710                  | 4.408                 | 4.781                  | 4.538                 | 4.440                  | 4.268                     |                         |
| V          | <sup>2</sup> Π <sub><i>g</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 4.949                  | 4.597                 | 5.079                  | 4.787                 | 4.613                  | 4.451                     | C:4.62(0.10)            |
| Ad         | <sup>2</sup> Π <sub><i>g</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.479                         | 180       | 4.812                  | 4.413                 | 4.895                  | 4.562                 | 4.509                  | 4.259                     |                         |
| V          | <sup>2</sup> Σ <sub><i>g</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.593                         | 180       | 4.883                  | 4.853                 | 4.982                  | 4.932                 | 4.577                  | 4.626                     | D:4.73(0.07)            |
| Ad         | <sup>2</sup> Σ <sub><i>g</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.582                         | 180       | 4.878                  | 4.860                 | 4.969                  | 4.938                 | 4.583                  | 4.643                     |                         |
| V          | <sup>2</sup> Δ <sub><i>g</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 5.758                  | 5.409                 | 5.895                  | 5.608                 | 5.373                  | 5.223                     | E:5.28(0.07)            |
| V          | <sup>2</sup> Δ <sub><i>u</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 6.001                  | 5.658                 | 6.140                  | 5.856                 | 5.551                  | 5.406                     | F:5.36(0.07)            |
| V          | <sup>2</sup> Π <sub><i>u</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 6.209                  | 5.889                 | 6.326                  | 6.061                 | 5.696                  | 5.548                     |                         |
| V          | <sup>2</sup> Σ <sub><i>u</i></sub> <sup>+</sup> (D <sub>∞<i>h</i></sub> ) | 2.593                         | 180       | 6.257                  | 5.930                 | 6.365                  | 6.091                 | 5.767                  | 5.601                     | G:5.53(0.05)            |
| V          | <sup>2</sup> Δ <sub><i>g</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 6.194                  | 5.866                 | 6.329                  | 6.060                 | 5.784                  | 5.650                     |                         |
| V          | <sup>2</sup> Π <sub><i>g</i></sub> (D <sub>∞<i>h</i></sub> )              | 2.593                         | 180       | 7.114                  | 6.828                 | 7.227                  | 6.995                 | 6.602                  | 6.483                     | H:5.90(0.05)            |

<sup>a</sup> Vertical IEs determined using the scalar relativistic CCSD/aug-cc-pVDZ-PP optimized Au<sub>3</sub><sup>-</sup> geometry, resulting in a D<sub>∞h</sub> structure with *r*<sub>Au-Au</sub> = 2.593 Å.

<sup>b</sup> Adiabatic IEs are the differences between the ground- and excited-state energies of Au<sub>3</sub> determined using the corresponding scalar relativistic IP-EOMCCSD(2*h*-1*p*)/aug-cc-pVDZ-PP optimized geometries and the ground-state energy of Au<sub>3</sub><sup>-</sup> determined as in footnote *a*.

<sup>c</sup> The 5d<sup>10</sup>6s<sup>1</sup> valence electrons of each gold atom and an additional electron due to the charge were correlated.

<sup>d</sup> The 5s<sup>2</sup>5p<sup>6</sup> semi-core and 5d<sup>10</sup>6s<sup>1</sup> valence electrons of each gold atom and an additional electron due to the charge were correlated.

<sup>e</sup> Extrapolated using Eq. (3.21).

<sup>f</sup> The labels, positions, and, in parentheses, widths of the peaks in the photoelectron spectrum of Au<sub>3</sub><sup>-</sup> from Ref. [222]. Except for C and E, which are shoulders of D and F, respectively, all peak positions correspond to the maxima in the spectrum. The width is the full width at half maximum (see Ref. [222] for details).



While all reported states are doublets, we did consider quartet states. However, the lowest-energy quartet, which was determined by experiment to lie in the energy range from 6.42–8.08 eV [236] and from the previously computed modified coupled pair functional (MCPF) ionization energies, which were artificially scaled, to be 7.06 eV [237], was found by our composite IP-EOMCC calculations to lie at 7.13 eV, i.e., in very good agreement with the previous experimental and theoretical studies and outside the range of the experimental photoelectron spectra [221–223]. All CC and IP-EOMCC calculations reported in this work were performed with the Piecuch group codes [159, 238] in GAMESS.

The IP-EOMCCSD( $2h-1p$ )/aug-cc-pVDZ-PP calculations for the first vertical IE of  $\text{Au}_3^-$  (peak A) are in good agreement with the experimental [222] value of 3.89 eV due to the fact that the ground state of  $\text{Au}_3$  is dominated by a  $1h$  excitation out of the  $\text{Au}_3^-$  HOMO (unlike in the early *ab initio* work [224–226], including high-level MRCI and MCPF computations, no empirical scaling had to be employed in our IP-EOMCC calculations to obtain such an agreement). The only other state of  $\text{Au}_3$  accurately described by this level of theory is the  $^2\Sigma_g^+$  state assigned to peak D, which is characterized by a leading  $1h$  excitation out of the HOMO – 1 orbital of  $\text{Au}_3^-$ , with all other excitations being much smaller. The remaining states have significant  $2h-1p$  contributions, resulting in errors of 0.27–1.21 eV relative to experiment at the IP-EOMCCSD( $2h-1p$ )/aug-cc-pVDZ-PP level. Looking at the IP-EOMCCSD( $2h-1p$ )/aug-cc-pVTZ-PP results, without and with CV contributions, all IEs rise on average by 0.11 and 0.16 eV, respectively, with the states assigned to peaks A and D changing the least. It is clear, comparing the IP-EOMCCSD( $2h-1p$ )/aug-cc-pV $x$ Z-PP ( $x = \text{D, T}$ ) and the corresponding IP-EOMCCSD( $2h-1p$ )/aug-cc-pV $x$ Z-PP+CV IEs, that the CV

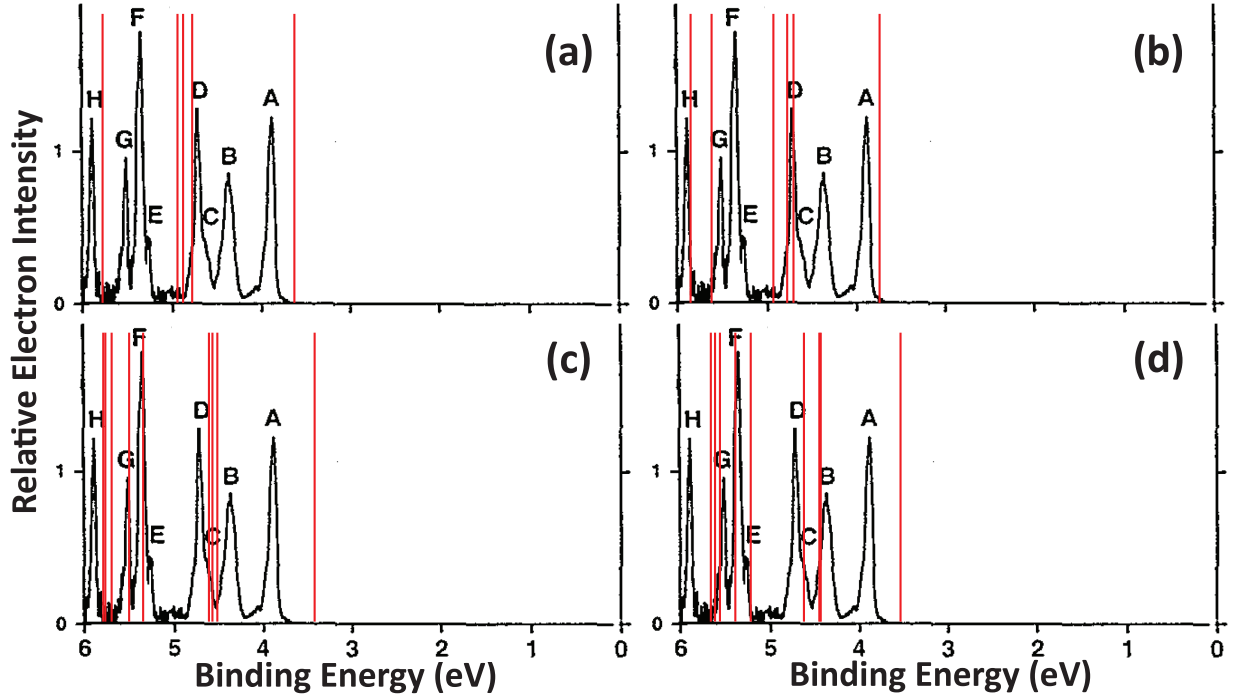


Figure 3.2: Vertical IEs of  $\text{Au}_3^-$  (red bars) superimposed on the photoelectron spectrum from Fig. 1(c) in Ref. [222]: (a) IP-EOMCCSD( $2h-1p$ )/aug-cc-pVDZ-PP calculations; (b) IP-EOMCCSD( $2h-1p$ )/aug-cc-pVTZ-PP+CV calculations; (c) IP-EOMCCSD( $3h-2p$ )/aug-cc-pVDZ-PP calculations; (d) extrapolated IEs determined using Eq. (3.21). For the actual symmetries and energies of the calculated states of  $\text{Au}_3$ , see Table 3.8.

correlations have a larger effect than the basis set employed, as they lower the calculated IEs by 0.23 eV on average, again having little effect on the states contributing to peaks A and D. The inclusion of  $3h-2p$  correlations within the IP-EOMCC framework has a significant effect on the majority of the calculated IEs, with an average lowering of  $\sim 0.4$  eV compared to the corresponding IP-EOMCCSD( $2h-1p$ )/aug-cc-pVDZ-PP calculations. Again, only the states assigned to peaks A and D, dominated by  $1h$  excitations, have smaller  $3h-2p$  correlations.

While the CV correlations and higher-order,  $3h-2p$ , excitations tend to lower the calculated IEs, the larger aug-cc-pVTZ-PP basis set raises them, thus demonstrating the need for a carefully balanced description of electron correlation and basis set effects when examining the electronic states of  $\text{Au}_3$ . Such a description would be obtained using the high-level IP-EOMCCSD( $3h-2p$ )/aug-cc-pVTZ-PP+CV approach, but calculations of this kind are too expensive in the  $\text{Au}_3^-$  case examined here. To circumvent this issue, we adopted the simple extrapolation scheme suggested in Section 3.3.2 which, as demonstrated in that section, can account for the CV as well as  $3h-2p$  correlations and the effect of going from the aug-cc-pVDZ-PP to the aug-cc-pVTZ-PP basis set in a computationally affordable manner by using Eq. (3.20) where BS1 is the aug-cc-pVDZ-PP basis set and BS2 the aug-cc-pVTZ-PP basis, defining the IE corresponding to the  $k$ -th electronic state of  $\text{Au}_3$ ,  $\text{IE}_k$ , as

$$\begin{aligned} \text{IE}_k = & [\text{IE}_k(3h-2p) - \text{IE}_k(2h-1p)]/\text{aug-cc-pVDZ-PP} \\ & + \text{IE}_k(2h-1p)/\text{aug-cc-pVTZ-PP+CV}. \end{aligned} \quad (3.21)$$

Similar to the explanation in Section 3.3.2, the first term on the right-hand side is the  $3h-2p$  contribution computed as the difference between the IP-EOMCCSD( $3h-2p$ ) and IP-

EOMCCSD( $2h-1p$ ) IEs using the aug-cc-pVDZ-PP basis, which is subsequently added to the IP-EOMCCSD( $2h-1p$ )/aug-cc-pVTZ-PP+CV IE accounting for the effects of the larger basis set and correlation of the semi-core electrons.

The vertical photoelectron spectrum obtained with Eq. (3.21) is shown in Fig. 3.2(d). With the exception of the ground state and our highest calculated excited state of  $\text{Au}_3$ , the extrapolated IEs are in excellent agreement with their experimental counterparts, deviating at most by 0.17 eV. Examining the ground state of  $\text{Au}_3$ , we can see that, while the inclusion of  $3h-2p$  excitations lowers the corresponding IE of  $\text{Au}_3^-$ , the larger aug-cc-pVTZ-PP basis set along with the CV correlations counteract this effect, resulting in a difference with experiment of 0.35 eV. Since the ground-state of  $\text{Au}_3$  is dominated by  $1h$  excitations relative to  $\text{Au}_3^-$ , one may have to use basis sets larger than aug-cc-pVTZ-PP in this case to further improve this result. The highest calculated excited state of  $\text{Au}_3$ , assigned to peak H, emphasizes the need to consider the various correlation effects together, as without all of them taken into account using Eq. (3.21) we obtain energies which are too high and which have not been accessed by the photodetachment experiments to date [221–223] (see Table 3.8).

In addition to an improved description of the IEs, our IP-EOMCC-based composite scheme defined by Eq. (3.21) allows one to accurately assign the underlying states of  $\text{Au}_3$  to peaks and regions of the photoelectron spectra of  $\text{Au}_3^-$  reported in Refs. [221–223], which would otherwise be a difficult thing to do (for the earlier attempts to examine excited states of  $\text{Au}_3$  at *ab initio* levels, see Refs. [227, 228]). Indeed, we computed three electronic states in the 4–5 eV region, all in good agreement with the experimental photodetachment spectra [222, 223]. The lowest two  $^2\Sigma_g^+$  states correspond to peaks B and D, while the lowest  $^2\Pi_g$

state can be assigned to the shoulder of peak D, labeled as peak C. For peaks E, F, and G our results indicate that there are five excited states of  $\text{Au}_3$  behind these spectral features. The lowest-energy  $^2\Delta_g$  state, which our extrapolation scheme based on Eq. (3.21) places at 5.223 eV, is clearly separated from the next  $^2\Delta_u$  state by 0.18 eV, and we assign it to the lowest-energy shoulder in this region, i.e., peak E. Of the remaining states in the same region, the two doubly degenerate  $^2\Delta_u$  and  $^2\Pi_u$  states, which our IP-EOMCC-based extrapolation scheme places at 5.406 and 5.548 eV, respectively, correspond to peak F, while the non-degenerate  $^2\Sigma_u^+$  and doubly degenerate  $^2\Delta_g$  states are assigned by us to peak G. This assignment is based on the relative intensities of the peaks observed in experiment, not the energies alone. Indeed, peak F has the largest intensity and, therefore, should be composed of more underlying electronic states than the other two peaks. Shoulder E, on the other hand, should originate from the least number of states, since it is the smallest peak in this region. Finally, the intensity of peak G lies between that of peaks E and F. Thus, there should be a, roughly, 2:4:3 ratio of the intensities between peaks E, F, and G, respectively, which is indeed the case if we use the above assignment and account for state degeneracies. It should be mentioned that this type of assignment for the intensities of the peaks, would be greatly enhanced if one computed the cross sections for each photodetached state, which would bring a more quantitative nature to the above elementary explanation of the peak heights.

In Ref. [222], the authors propose that in the higher-energy region above 5 eV  $\text{Au}_3$  possesses an  $s$ -derived  $^2\Sigma_g^+$  state, meaning that the electron ejected from  $\text{Au}_3^-$  originates from an orbital with a dominant  $s$  character. They use empirical arguments deduced from

the photoelectron spectra of group IB elements (Cu, Ag, Au) and assign the  $s$ -derived  $^2\Sigma_g^+$  state to peak F. Examining our IP-EOMCC wave functions, we can see that the photoelectron spectrum of  $\text{Au}_3^-$  indeed contains an  $s$ -derived  $^2\Sigma_g^+$  state. However, our IP-EOMCC-based extrapolation scheme using Eq. (3.21) places such a state at 4.626 eV, if vertical IEs are examined, and so we assign it to peak D. The  $^2\Sigma_u^+$  ground state of  $\text{Au}_3$  is the only other  $s$ -derived state, with the remaining states being  $d$ -derived, in agreement with the discussion in Ref. [222]. In the presence of spin-orbit coupling, ignored in this work, the manifold of the  $d$ -derived states would further split into the  $d_{5/2}$ - and  $d_{3/2}$ -derived bands [222], but based on the earlier computations [227] we do not expect this to alter the physical origin of the peaks and shoulders in the experimental photoelectron spectrum suggested by our scalar relativistic IP-EOMCC calculations.

While our extrapolated vertical IEs of  $\text{Au}_3$  are in generally good agreement with experiment, using them alone is not sufficient to explain the varying peak widths observed in the photoelectron spectrum, especially in the lower-energy region, 3–5 eV. It is possible, and has been observed in other clusters [222, 223, 239, 240], that ejection of an electron can be a slow enough process to allow for the geometry of the cluster to relax, contributing to the experimental peak widths. For this reason, we investigated the role of geometry relaxation on the ground and lower-energy (<5 eV) excited states of  $\text{Au}_3$ , using the IP-EOMCCSD( $2h$ - $1p$ )/aug-cc-pVDZ-PP level of theory and finite-difference approach to optimize their geometries. All of the geometries were optimized using  $C_s$  symmetry employing as the starting geometry a bent structure with one of the Au–Au bond lengths longer than the other Au–Au distance. We again employed our extrapolation scheme based on Eq. (3.21) to

determine the final energetics.

For the ground state of  $\text{Au}_3$ , two low-lying isomers ( $X_0$  and  $X_1$  in Table 3.8) were found. They are consistent with previous theoretical [208, 210–212, 224, 227] and experimental [241] work. The lowest-energy isomer ( $X_0$ ) is an isosceles, nearly  $D_{3h}$  symmetric, Jahn-Teller distorted triangle (the spin-orbit effects, ignored in this work, would suppress this distortion [210, 212]). Transition from the linear geometry of  $\text{Au}_3^-$  to the triangular  $X_0$  geometry of  $\text{Au}_3$  requires significant lengthening of the Au–Au bonds and a large change of the Au–Au–Au angle,  $\theta$ , i.e., isomer  $X_0$  may be hard to access in photoelectron experiments. The  $X_1$  isomer (also found in Ref. [210]), which is characterized by the nearest-neighbor Au–Au distance similar to that in  $\text{Au}_3^-$  and the Au–Au–Au angle  $\theta = 142^\circ$ , seems easier to access from the equilibrium geometry of  $\text{Au}_3^-$ . The 0.087 eV difference between the extrapolated adiabatic IE for isomer  $X_1$  and the corresponding vertical IE agrees to within 0.01 eV with the experimentally reported full width at half maximum (FWHM) characterizing peak A. Peak B is nearly twice as broad as peak A in experiment, agreeing with our computed peak widths for A and B of 0.087 and 0.180 eV, respectively. Since peaks C and D are not resolved in the experiment where their FWHM was measured [222] (cf. Ref. [223] where peak C and D are resolved in the experimental spectrum, though their peak positions and FWHMs are not determined there) and the total width of their combined feature is nearly the same as the difference between our calculated vertical and adiabatic IEs for the lowest  $^2\Pi_g$  and second-lowest  $^2\Sigma_g^+$  states, their analysis is not as straightforward. Nevertheless, the difference between the vertical and adiabatic IEs obtained using our IP-EOMCC extrapolation for the lowest-energy  $^2\Pi_g$  state assigned to shoulder C is larger than that characterizing the second

$^2\Sigma_g^+$  state assigned to peak D, in qualitative agreement with the experimental peak widths. As for the higher-lying states above 5 eV, the difference between vertical IEs characterizing the  $^2\Delta_u$  and  $^2\Pi_u$  states, assigned to peak F, and  $^2\Sigma_u^+$  and  $^2\Delta_g$  states, assigned to peak G, are comparable to the corresponding experimental peak widths, strengthening our interpretation that each of these two spectral features originates from multiple electronic states. The high density of states in this region made the geometry optimizations uncertain, so we do not discuss the corresponding adiabatic IEs.

In summary, we investigated the photoelectron spectrum of  $\text{Au}_3^-$  using the scalar relativistic IP-EOMCCSD( $2h-1p$ ) and IP-EOMCCSD( $3h-2p$ ) approaches. We examined the effects of basis set, number of correlated electrons, higher-order correlation contributions, and geometry relaxation, obtaining an accurate assignment of peaks and shoulders observed in the experimental photoelectron spectrum. The final energetics were obtained with an IP-EOMCC-based composite scheme capable of capturing the basis set, core-valence, and higher-order valence correlation effects in a computationally feasible manner, producing the results that on average agree with the experimental peaks to within  $\sim 0.1$  eV. Our calculations suggest that changes in geometry during electron ejection from  $\text{Au}_3^-$  may contribute to the peak widths, in addition to multiple electronic states behind a given spectral feature.

In future studies, we would like to determine the effects of higher-order relativistic contributions, such as spin-orbit interactions, on our results as well as test the reliability of the IP-EOMCC approaches with  $3h-2p$  and  $4h-3p$  excitations, treated fully or with active orbitals [160] and extrapolation schemes of the type of Eqs. (3.20) and (3.21). It would also be interesting to include scattering cross-sections in our future analysis.



# Chapter 4

## Applications of completely renormalized coupled-cluster and equation-of-motion coupled-cluster approaches

In this chapter, the completely renormalized (CR) CC ground-state and CR-EOMCC excited-state methods are used to investigate several challenging chemical problems, from the photochemistry of organic molecules to the catalytic reaction pathways involving gold nanoparticles. Each of our examples demonstrates the utility of novel single-reference CC and EOMCC approaches developed in our group in providing accurate information and predictive power.

### 4.1 Theory

In the basic particle-conserving CCSD and EOMCCSD approaches, the cluster operator  $T$  and the linear excitation operator  $R_\mu$  defining the corresponding ground-state and excited-state wave functions,  $|\Psi_0\rangle = e^T|\Phi\rangle$  and  $|\Psi_\mu\rangle = R_\mu|\Psi_0\rangle = R_\mu e^T|\Phi\rangle$ , respectively, are trun-

cated at two-body terms, such that (cf. Eq. (3.11))  $T \approx T^{(\text{CCSD})} = T_1 + T_2$  where

$$T_1 = \sum_{i,a} t_a^i a^a a_i \quad (4.1)$$

and

$$T_2 = \sum_{i < j, a < b} t_{ab}^{ij} a^a a^b a_j a_i, \quad (4.2)$$

and  $R_\mu \approx R_\mu^{(\text{CCSD})} = R_{\mu,0} + R_{\mu,1} + R_{\mu,2}$ , where

$$R_{\mu,0} = r_{\mu,0} \mathbf{1} \quad (4.3)$$

with  $\mathbf{1}$  representing the unit operator, and

$$R_{\mu,1} = \sum_{i,a} r_a^i a^a a_i \quad (4.4)$$

and

$$R_{\mu,2} = \sum_{i < j, a < b} r_{ab}^{ij} a^a a^b a_j a_i. \quad (4.5)$$

As before, we adopt the conventional notation in which  $i, j, \dots$  ( $a, b, \dots$ ) designate the occupied (unoccupied) spin-orbitals in the reference determinant  $|\Phi\rangle$  and  $a^p$  ( $a_p$ ) represent the usual creation (annihilation) operators associated with the spin-orbital basis set  $\{|p\rangle\}$ .

The singly and doubly excited cluster amplitudes,  $t_a^i$  and  $t_{ab}^{ij}$ , defining  $T_1$  and  $T_2$ , respectively, are obtained by solving the system of nonlinear equations obtained by projecting the electronic Schrödinger equation with the CCSD wave function in it onto the singly and

doubly excited determinants,  $|\Phi_i^a\rangle = a^a a_i |\Phi\rangle$  and  $|\Phi_{ij}^{ab}\rangle = a^a a^b a_j a_i |\Phi\rangle$ , namely,

$$\langle \Phi_i^a | \bar{H}^{(\text{CCSD})} | \Phi \rangle = \langle \Phi_i^a | \bar{H}_{N,\text{open}}^{(\text{CCSD})} | \Phi \rangle = 0 \quad (4.6)$$

and

$$\langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle = \langle \Phi_{ij}^{ab} | \bar{H}_{N,\text{open}}^{(\text{CCSD})} | \Phi \rangle = 0, \quad (4.7)$$

where (cf. Eqs. (3.10) and (3.13))

$$\bar{H}^{(\text{CCSD})} = e^{-T_1 - T_2} H e^{T_1 + T_2} = \bar{H}_{N,\text{open}}^{(\text{CCSD})} + E_0^{(\text{CCSD})} \quad (4.8)$$

is the similarity transformed Hamiltonian of CCSD. The ground-state CCSD energy,  $E_0^{(\text{CCSD})}$ , is obtained by projecting the Schrödinger equation onto the reference determinant  $|\Phi\rangle$ ,

$$E_0^{(\text{CCSD})} = \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle = \langle \Phi | H | \Phi \rangle + \bar{H}_{N,\text{closed}}^{(\text{CCSD})}. \quad (4.9)$$

In the EOMCCSD case, the one- and two-body excitation amplitudes,  $r_a^i$  and  $r_{ab}^{ij}$ , defining  $R_{\mu,1}$  and  $R_{\mu,2}$ , respectively, and the corresponding vertical excitation energies,

$$\omega_\mu^{(\text{CCSD})} = E_\mu^{(\text{CCSD})} - E_0^{(\text{CCSD})}, \quad (4.10)$$

are obtained by diagonalizing the similarity transformed Hamiltonian  $\bar{H}^{(\text{CCSD})}$  (or  $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$ ) in the space spanned by the singly and doubly excited determinants,  $|\Phi_i^a\rangle$  and  $|\Phi_{ij}^{ab}\rangle$ , respec-

tively,

$$\langle \Phi_i^a | (\bar{H}_{\text{open}}^{(\text{CCSD})} R_{\mu, \text{open}}^{(\text{CCSD})})_C | \Phi \rangle = \omega_{\mu}^{(\text{CCSD})} r_a^i \quad (4.11)$$

and

$$\langle \Phi_{ij}^{ab} | (\bar{H}_{\text{open}}^{(\text{CCSD})} R_{\mu, \text{open}}^{(\text{CCSD})})_C | \Phi \rangle = \omega_{\mu}^{(\text{CCSD})} r_{ab}^{ij}, \quad (4.12)$$

combined with

$$r_{\mu, 0} = \langle \Phi | (\bar{H}_{\text{open}}^{(\text{CCSD})} R_{\mu, \text{open}}^{(\text{CCSD})})_C | \Phi \rangle / \omega_{\mu}^{(\text{CCSD})} \quad (4.13)$$

for the zero-body component or  $R_{\mu}$ , which is computed *a posteriori*.

Since the left (bra) and right (ket) eigenstates of the non-Hermitian similarity transformed Hamiltonian  $\bar{H}^{(\text{CCSD})}$  form a biorthogonal many-electron basis, they must both be determined for evaluation of properties other than energies or if one is interested in the subsequent CR-CC and CR-EOMCC computations. The proper ansatz for the bra wave functions of the CC/EOMCC theory is  $\langle \tilde{\Psi}_{\mu} | = \langle \Phi | L_{\mu} e^{-T}$ , where  $\mu = 0$  indicates the ground state and  $\mu > 0$  labels excited states, which in the CCSD/EOMCCSD case are generated by the deexcitation operators  $L_{\mu} \approx L_{\mu}^{(\text{CCSD})} = L_{\mu, 0} + L_{\mu, 1} + L_{\mu, 2}$ , where

$$L_{\mu, 0} = \delta_{\mu, 0} \mathbf{1}, \quad (4.14)$$

$$L_{\mu, 1} = \sum_{i, a} l_i^a a^i a_a, \quad (4.15)$$

and

$$L_{\mu, 2} = \sum_{i < j, a < b} l_{ij}^{ab} a^i a^j a_b a_a, \quad (4.16)$$

whose many-body ranks match those of  $T^{(\text{CCSD})}$  and  $R_\mu^{(\text{CCSD})}$ .

The noniterative triples corrections to the CCSD and EOMCCSD energies defining the CR-CC(2,3) and CR-EOMCC(2,3) approaches result from the more general theoretical framework, abbreviated as MMCC [46–52, 87–95] (cf. Ref. [40, 41, 68] for selected reviews). Thus, they are based on one of the several possible expressions [40, 41, 46–52, 68, 87–95] for the differences between the exact, full CI energies  $E_\mu$  and the corresponding CC/EOMCC energies  $E_\mu^{(\text{CC})}$  obtained with truncated forms of  $T$  and  $R_\mu$ ,

$$\delta_\mu = E_\mu - E_\mu^{(\text{CC})}, \quad (4.17)$$

written in terms of the generalized moments of the CC/EOMCC equations corresponding to *a posteriori* projections of these equations on the excited determinants disregarded in the CC/EOMCC calculations. In the specific case of CR-CC(2,3) and CR-EOMCC(2,3) considerations, we write the corresponding energies  $E_\mu(2, 3)$  as follows:

$$E_\mu(2, 3) = E_\mu^{(\text{CCSD})} + \delta_\mu(2, 3), \quad (4.18)$$

where the triples corrections  $\delta_\mu(2, 3)$  to the CCSD ( $\mu = 0$ ) and EOMCCSD ( $\mu > 0$ ) energies  $E_\mu^{(\text{CCSD})}$  are calculated as

$$\delta_\mu(2, 3) = \sum_{i < j < k, a < b < c} \ell_{\mu,ijk}^{abc} \mathfrak{M}_{\mu,abc}^{ijk}. \quad (4.19)$$

The generalized moments of the CCSD and EOMCCSD equations  $\mathfrak{M}_{\mu,abc}^{ijk}$  entering Eq. (4.19)

are obtained by projecting the CCSD and EOMCCSD equations on the triply excited determinants  $|\Phi_{ijk}^{abc}\rangle$ ,

$$\mathfrak{M}_{\mu,abc}^{ijk} = \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} R_{\mu}^{(\text{CCSD})} | \Phi \rangle, \quad (4.20)$$

where  $R_{\mu}^{(\text{CCSD})} = \mathbf{1}$  when  $\mu = 0$ , the CCSD ground-state case is considered. The corresponding amplitudes  $\ell_{\mu,ijk}^{abc}$  that multiply moments  $\mathfrak{M}_{\mu,abc}^{ijk}$  to produce the  $\delta_{\mu}(2,3)$  correction are calculated using

$$\ell_{\mu,ijk}^{abc} = \langle \Phi | (L_{\mu,0} + L_{\mu,1} + L_{\mu,2}) \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle / D_{\mu,ijk}^{abc}, \quad (4.21)$$

where  $D_{\mu,ijk}^{abc}$  is a suitable quasi-perturbative denominator derived by considering the left eigenvalue problem involving  $\bar{H}^{(\text{CCSD})}$  in a subspace spanned by up to triple excitations. Different forms of  $D_{\mu,ijk}^{abc}$  lead to different variants of the CR-CC(2,3) and CR-EOMCC(2,3) approaches labeled by an additional letter A–D. In particular, in the most complete variant D resulting in the CR-CC(2,3),D and CR-EOMCC(2,3),D methods,  $D_{\mu,ijk}^{abc}$  entering Eq. (4.21) is calculated using the Epstein-Nesbet-like expression,

$$D_{\mu,ijk}^{abc} = \omega_{\mu}^{(\text{CCSD})} - \sum_{n=1}^3 \langle \Phi_{ijk}^{abc} | \bar{H}_n^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle, \quad (4.22)$$

where  $\bar{H}_n^{(\text{CCSD})}$  is the  $n$ -body component of  $\bar{H}^{(\text{CCSD})}$ . On the other hand, in the simplest A version of CR-CC(2,3) and CR-EOMCC(2,3) which leads to the CR-CC(2,3),A and CR-EOMCC(2,3),A approaches, we replace the Epstein-Nesbet form of  $D_{\mu,ijk}^{abc}$ , Eq. (4.22), by

the Møller-Plesset-style expression

$$D_{\mu,ijk}^{abc} = \omega_{\mu}^{(\text{CCSD})} - (\varepsilon_a + \varepsilon_b + \varepsilon_c - \varepsilon_i - \varepsilon_j - \varepsilon_k), \quad (4.23)$$

where  $\varepsilon$ 's are the Hartree-Fock spin-orbital energies (diagonal elements of the Fock matrix). As explained in Refs. [93–95], the ground-state CR-CC(2,3)<sub>A</sub> scheme is equivalent to the CCSD(2)<sub>T</sub> approach of Ref. [77]. Similarly, the CR-EOMCC(2,3)<sub>A</sub> method is equivalent to the EOM-CC(2)PT(2) approach of Ref. [76].

One of the key features of the CC and EOMCC theories is size extensivity of the CC results in the ground state and size intensivity of the EOMCC results in describing excitation energies. The ground-state CR-CC(2,3) approach is size extensive, so there is no concern here. Unfortunately, the CR-EOMCC(2,3) triples corrections to the EOMCCSD excited-state energies (just like their EOM-CC(2)PT(2) counterpart) introduce small size-intensivity errors (the EOMCCSD approach is size intensive). In order to address this issue, we have to turn to the  $\delta$ -CR-EOMCC(2,3) methodology developed in Ref. [31] as a consequence of the analysis presented in Refs. [45] and [52].

The  $\delta$ -CR-EOMCC(2,3) method proposed in Ref. [31], which is a rigorously size-intensive version of the CR-EOMCC(2,3) methodology of Refs. [51, 52], relies on the same general expression for the triples corrections  $\delta_{\mu}(2, 3)$ , but differs in the explicit equations for  $\mathfrak{M}_{\mu,abc}^{ijk}$  and in the fact that the EOMCCSD excitation energies  $\omega_{\mu}^{(\text{CCSD})}$  rather than the total energies  $E_{\mu}^{(\text{CCSD})}$  are corrected for the missing triples contributions,

$$\omega_{\mu}(2, 3) = \omega_{\mu}^{(\text{CCSD})} + \delta_{\mu}(2, 3), \quad (4.24)$$

where  $\mu > 0$ . As shown in Refs. [31, 52], the enforcement of strict size intensity of the  $\delta_\mu(2, 3)$  triples corrections to the EOMCCSD excitation energies ( $\omega_\mu^{(\text{CCSD})}$  is size intensive [28, 242]) requires that we replace the complete moment  $\mathfrak{M}_{\mu,abc}^{ijk}$ , Eq. (4.20), in Eq. (4.19) for  $\delta_\mu(2, 3)$  by its truncated analog ignoring the ground-state  $r_{\mu,0}\langle\Phi_{ijk}^{abc}|\bar{H}^{(\text{CCSD})}|\Phi\rangle$  contribution, i.e.,

$$\mathfrak{M}_{\mu,abc}^{ijk} = \langle\Phi_{ijk}^{abc}|\bar{H}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})|\Phi\rangle. \quad (4.25)$$

We continue using Eq. (4.21) for the deexcitation amplitudes  $\ell_{\mu,ijk}^{abc}$ , which in the case of excited states targeted by  $\delta$ -CR-EOMCC(2,3) are calculated as (cf. Eq. (4.14))

$$\ell_{\mu,ijk}^{abc} = \langle\Phi|(L_{\mu,1} + L_{\mu,2})\bar{H}^{(\text{CCSD})}|\Phi_{ijk}^{abc}\rangle/D_{\mu,ijk}^{abc}. \quad (4.26)$$

Again, if the  $D_{\mu,ijk}^{abc}$  denominator entering Eq. (4.26) is given by the Epstein-Nesbet-type expression, Eq. (4.22), we obtain the more complete variant D of  $\delta$ -CR-EOMCC(2,3), abbreviated as  $\delta$ -CR-EOMCC(2,3),D. If we replace Eq. (4.22) for  $D_{\mu,ijk}^{abc}$  in Eq. (4.26) by the Møller-Plesset-type expression given by Eq. (4.23), we obtain the simplified A variant, abbreviated as  $\delta$ -CR-EOMCC(2,3),A, which is, as explained in Ref. [31] (cf., also, Ref. [52]), equivalent to the EOMCCSD(2)<sub>T</sub> approach of Ref. [45] and, if we limit ourselves to vertical excitation energies, to the EOMCCSD( $\tilde{T}$ ) method of Ref. [86]. Typically, we use the  $\delta$ -CR-EOMCC(2,3) approach to calculate excitation energies. If we are, however, interested in total  $\delta$ -CR-EOMCC(2,3) energies, we simply add  $\omega_\mu(2, 3)$ , Eq. (4.24), to the ground-state CR-CC(2,3) energy. As a result, the  $\delta$ -CR-EOMCC(2,3) method is extensive in the ground state (using the CR-CC(2,3) expressions for the ground-state energy) and size intensive in



describing excitation energies, satisfying the key desiderata of the CC/EOMCC theories.

## 4.2 Discovery of the doubly excited state that mediates the photoionization of azulene

### 4.2.1 Background information and performed calculations

Typically, when talking about molecular electronic spectroscopy, the focus is on dipole-allowed transitions to states with predominant one-electron excitation character, which make up the majority of photoabsorption spectra. It is well known, though, that molecular systems can possess strongly correlated excited states dominated by two-electron transitions. Such states, while being spectroscopically dark when populated from the ground states, can be useful tools for probing wider ranges of vibrational energies and novel types of photoinduced chemical dynamics. As a part of this thesis work, we examined the electronic spectrum of azulene combining high-level *ab initio* quantum chemistry with experiment to prove the existence of a doubly excited state below the ionization threshold which can drive the intriguing multiphoton ionization dynamics examined in Ref. [243], resulting in clear Rydberg fingerprint spectra.

The electronic structure of azulene, whose molecular configuration is shown in Fig. 4.1, is noteworthy due to its atypical fluorescence, which occurs from the second-excited  $S_2$  state, instead of the lower-energy  $S_1$  state. As such, azulene is a textbook exception to Kasha’s rule [244], which, although originally formulated for condensed phase matter, applies to many gas phase molecules as well, says that “the emitting level of a given multiplicity is

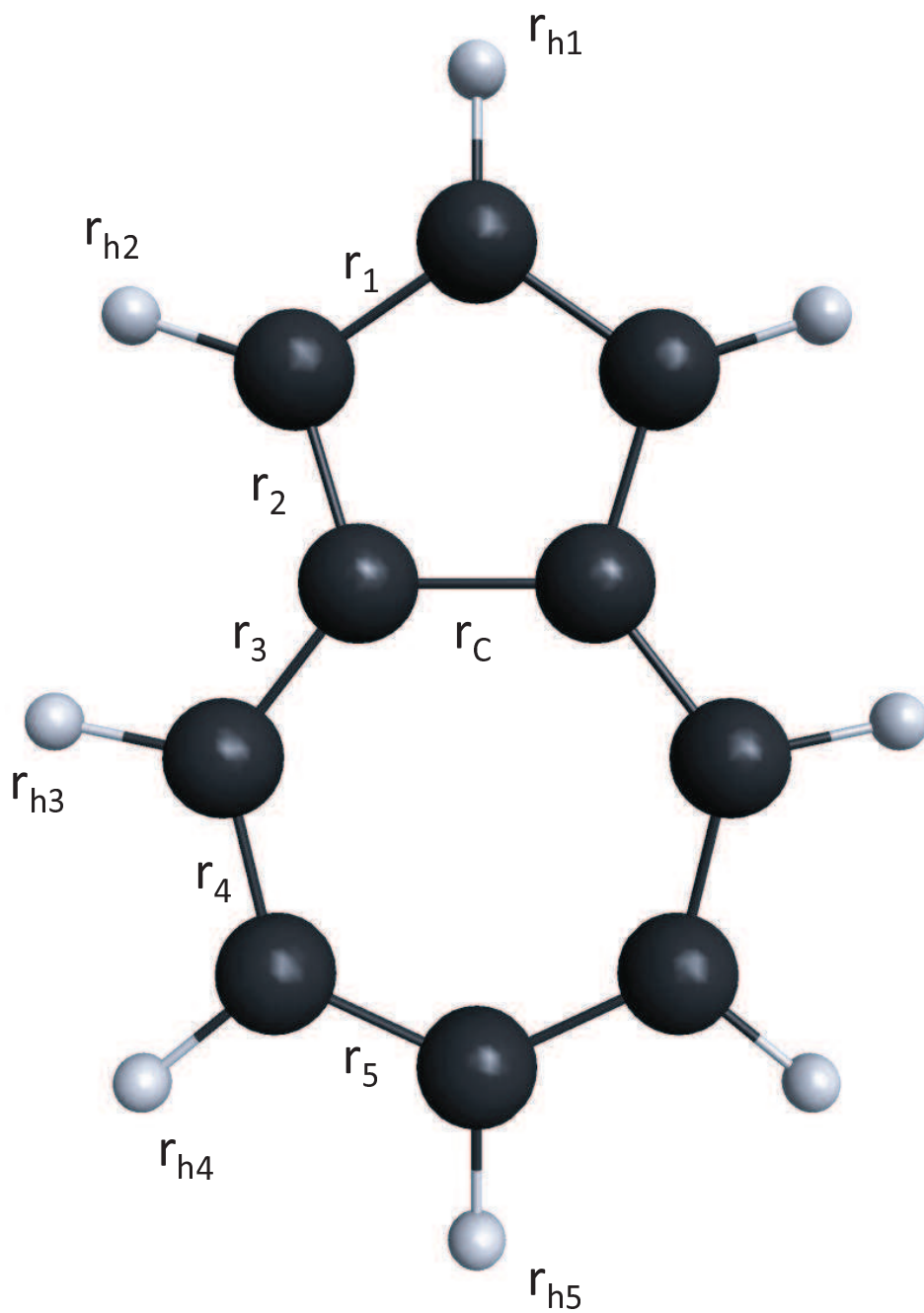


Figure 4.1: The  $C_{2v}$ -symmetric structure of azulene and its bond lengths. The large black spheres represent carbon atoms, whereas the small gray spheres represent hydrogens.

the lowest excited level of that multiplicity.” Due to this unusual behavior, there have been many studies, both experimental and theoretical, on the photochemistry of azulene, including several experimental studies dealing with the spectroscopy and dynamics involving the  $S_1$  and  $S_2$  states [245–248]. The early two-photon ionization experiments via the  $S_n$ ,  $n = 2 - 4$ , states of azulene [249] should be mentioned here, too. Much less is known, however, about higher valence states of azulene [250–253], other than their utility in preparing highly excited molecules for observation of photoinduced unimolecular dynamics [254–256].

Recently, Blanchet and coworkers carried out an extensive experimental study of the photoionization of azulene [243]. The photoelectron spectra of azulene recorded in their two-color, three-photon,  $1 + 2'$  time-resolved photoelectron imaging experiments are shown in Fig. 4.2, which combines the photoelectron spectra from Ref. [243], where the pump wavelengths were varied from 268 to 335 nm and probe pulses were fixed at 400nm, with the analogous spectrum corresponding to the 201 nm pump obtained in Ref. [142] (see Ref. [257] for a related study). They noticed that the  $1 + 2'$  time-resolved photoelectron spectra are invariant (apart from the intensity) with respect to the pump-probe delay time and wavelength of the pump pulse (cf. Fig. 4.2), which led them to suggest that the  $1 + 2'$  photoionization is driven by an unstable doubly excited electronic state located below the ionization threshold [243]. In the study carried out in Ref [243] Blanchet and coworkers stated “...we hope our observations might inspire theorists to take up the challenge to calculate the geometry and electronic configuration of the doubly excited states on polyaromatic systems, such as azulene”, as they had no direct way to prove experimentally that such a state existed. An important note about the postulated doubly excited state in the middle of the Rydberg

part of azulene’s electronic spectrum, is that it should be distinguished from “superexcited” resonance states above the ionization threshold that the normal, basis-set-based, real space *ab initio* quantum chemistry approaches for electronically bound states, including those employed in this dissertation, cannot describe. As depicted in Fig. 4.3, the postulated doubly excited state, marked by \*\*, is proposed to be populated by a probe transition from the  $S_2$  state, which itself is populated by the initial pump photon or by fast relaxation from the  $S_n$  states with  $n > 2$  reached by the pump energies. Once populated, the doubly excited state is proposed to rapidly relax into the vibrationally highly excited Rydberg states, from which azulene can be photoionized after absorbing the second probe photon to produce a Rydberg fingerprint.

The available experimental information suggests that the postulated doubly excited state is located not only below the ionization threshold of 7.41 eV, but it must, in fact, appear below 6.81 eV (a combination of the lowest-energy pump photon at 335 nm and a 400 nm probe photon). It should also be located above the second Rydberg state at 5.19 eV, marked in Fig. 4.2 as  $R_B$  (cf. Table 4.2), since the fingerprint of this state is clearly seen in the lower-energy parts of all the photoelectron spectra shown in Fig. 4.2, independent of the pump wavelength. Furthermore, to rationalize the optically induced anisotropy associated with the pump excitation [257], the postulated doubly excited state should represent a totally symmetric singlet excitation accessible from  $S_2$ . Deleuze carried an exhaustive investigation using algebraic-diagrammatic construction calculations at the ADC(3) level of theory for azulene and other polycyclic aromatic hydrocarbons and their ionization spectra [258]. He showed that the first shake-up ion state of the azulene cation, which lies at very low energy, consists of a dominant orbital configuration of  $\dots(3b_1)^2(2a_2)^0(4b_1)^0(3a_2)^1$  with respect to the neutral ground-state electron configuration,  $\dots(3b_1)^2(2a_2)^2(4b_1)^0(3a_2)^0$ . Since the postulated doubly excited state lies nears the ionization level and the Rydberg states are assumed to converge to this cation state, the postulated doubly excited state should have a similar electron configuration dominated by the  $(\text{HOMO})^2 \rightarrow (\text{LUMO} + 1)^2$  and  $(\text{HOMO} - 1)^2 \rightarrow (\text{LUMO})^2$  transitions [243]. It should be mentioned that the multiphoton ionization of azulene from  $S_2$  via unstable superexcited valence states relaxing to the Rydberg states prior to the final photoionization event has also been discussed in Ref. [259], but without determining a precise makeup of the hypothetical superexcited states. Moreover, superexcited states are located above the ionization threshold, whereas the doubly excited

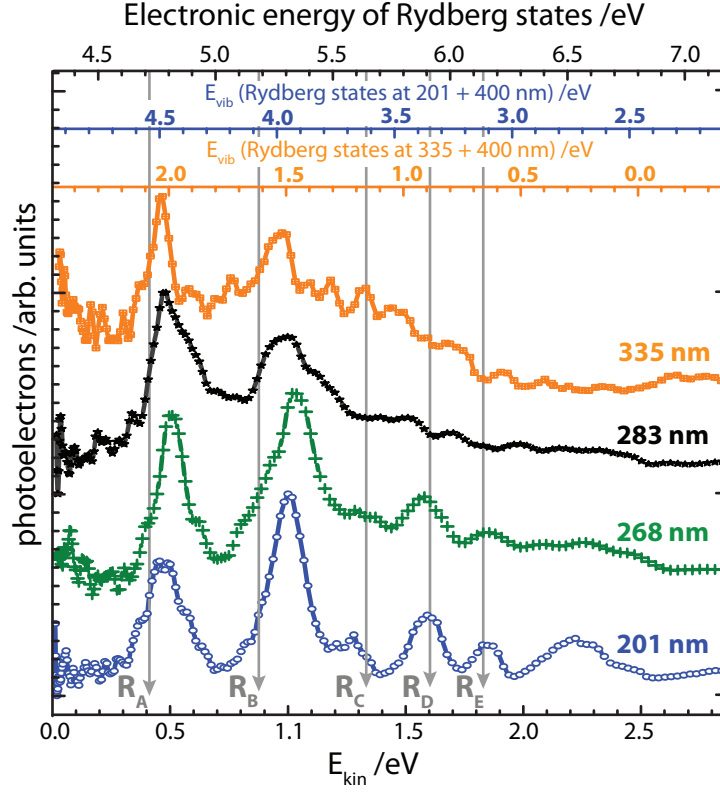


Figure 4.2: The photoelectron spectrum of azulene using a probe pulse centered at 400 nm, recorded for four pump excitation energies (the corresponding vibrational energies in  $S_2$  given in parentheses), namely, 201 nm (2.61 eV; Ref. [142]), 268 nm (1.07 eV; Ref. [243]), 283 nm (0.82 eV; Ref. [243]), and 335 nm (0.14 eV; Ref. [243]), and a pump-probe delay time of 500 fs (for other pump-probe delay times, which yield similar spectra, see Ref. [243]). The unstructured direct ionization (65 % of photoionization events) has been subtracted. Each spectral profile represents a fingerprint of the Rydberg states from which azulene is photoionized after absorbing the second probe photon. The electronic energies of these Rydberg states are marked by  $R_A$ ,  $R_B$ , etc. Two horizontal axes near the top show the vibrational energies in Rydberg states for the pump wavelengths of 201 and 335 nm.

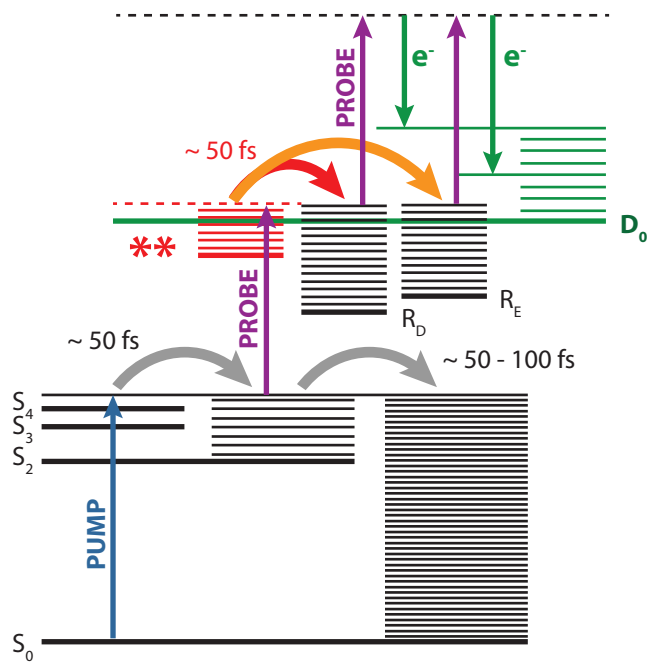


Figure 4.3: The proposed schematics of the  $1 + 2'$  photoionization experiment [243]. The Rydberg states, from which azulene is photoionized, are populated by the electronic relaxation from the postulated doubly excited state located below the ionization threshold  $D_0$ , marked by **\*\***. Excitation of the doubly excited state from  $S_2$ , relaxation into the Rydberg states, and photoionization take place within the probe pulse duration (100 fs).

state postulated in Ref. [243] and sought in this thesis research is expected to lie below the ionization threshold. We note that doubly excited states at energies as low as 4.95 eV have been detected by the magnetic circular dichroism experiments on azulene derivatives [260].

To determine if the proposed doubly excited state indeed exists, we carried out a large number of  $\delta$ -CR-EOMCC(2,3) calculations, including low-lying valence states and several states in the higher-energy region near the ionization threshold [142]. As explained in Section 4.1, the  $\delta$ -CR-EOMCC(2,3) approach is the rigorously size-intensive modification of the CR-EOMCC(2,3) method which is, in turn, the extension of the CR-CC(2,3) approach excited electronic states. Just like CR-EOMCC(2,3) and its predecessors [48, 49], or methods mentioned in Section 4.1, such as EOMCCSD( $\tilde{T}$ ) and EOMCCSD(2)<sub>T</sub>, the  $\delta$ -CR-EOMCC(2,3) approach corrects the vertical excitation energies obtained with the EOMCCSD scheme for the effects of triple excitations that are necessary to accurately describe excited states dominated by two-electron transitions within the EOMCC framework. As mentioned in the Introduction, EOMCCSD describes the energetics of excited states dominated by two-electron transitions poorly, pushing them to much higher energies (cf., e.g., Refs. [18, 40–43, 46–52, 54, 85, 86]). Methods, such as EOMCCSD( $\tilde{T}$ ), EOMCCSD(2)<sub>T</sub>, and especially  $\delta$ -CR-EOMCC(2,3) provide the necessary energy lowering. The  $\delta$ -CR-EOMCC(2,3) approach has a few variants, discussed, in particular, in Section 4.1, including the more complete variant D and the simplified variant A equivalent to EOMCCSD(2)<sub>T</sub> and EOMCCSD( $\tilde{T}$ ), but we only show variant D calculations, since the results obtained with variant A are rather similar and do not alter our main conclusions.

Our  $\delta$ -CR-EOMCC(2,3) calculations for azulene were performed using the 6-31G(d) and



cc-pVDZ basis sets, which were the largest we could reasonably accommodate when calculating so many excited states of azulene at such a high EOMCC level. The initial  $\delta$ -CR-EOMCC(2,3)/6-31G(d) calculations were performed at the ground-state geometry optimized at the MP2/6-31G(d) level with GAMESS, where the  $\delta$ -CR-EOMCC(2,3) and other EOMCC routines developed by the Piecuch group [31, 48, 49, 51, 52, 145] are incorporated. The final  $\delta$ -CR-EOMCC(2,3)/cc-pVDZ calculations were performed at the improved ground-state geometry obtained using the numerical CR-CC(2,3)/cc-pVDZ gradients in GAMESS, consistent with the  $\delta$ -CR-EOMCC(2,3)/cc-pVDZ description of excited states, which we verified with our parallel numerical derivatives discussed in Section 5.1. All calculations used the RHF reference and the 10 lowest-energy core orbitals corresponding to the 1s shells of the carbon atoms were frozen in the post-SCF considerations. Although basis sets used here cannot describe the Rydberg states, they are acceptable for representing one- and two-electron valence transitions, including the doubly excited state we have attempted to find, as long as one uses a higher-level EOMCC methodology with a robust treatment of triple excitations, such as  $\delta$ -CR-EOMCC(2,3) (in the Franck-Condon region of the pump and probe transitions, the interaction with the Rydberg states is not expected to strongly perturb the vertical excitation energy of the doubly excited valence state).

### 4.2.2 Results

The results of our MP2/6-31G(d) and CR-CC(2,3)/cc-pVDZ geometry optimizations performed prior to the final  $\delta$ -CR-EOMCC(2,3) work, are shown in Table 4.1 along with the available experimental and previously obtained theoretical data. Our equilibrium geome-

tries are in good agreement with the other theoretical [261–264] and experimentally determined [265–268] results. Table 4.2 shows the ten lowest excited states we found along with the available multireference complete active-space self consistent field (CASSCF) [269, 270], CASSCF based second-order perturbation theory (CASPT2) [271, 272], and experimental vertical excitation energies [243]. The number of electronic states we had to converge had to be much larger than in the case of the earlier CASSCF and CASPT2 calculations of the low-lying excitations in azulene dominated by one-electron transitions [262], which shows the advantage of using our single-reference EOMCC ideas over the multireference CASSCF-based thinking, since choosing an adequate active-space to encompass all of the singlet excited states of azulene and considered in this work would be problematic and prohibitively expensive. Additional information about the dominant excitation amplitudes defining the EOMCCSD wave functions for the calculated excited states is given in Table 4.3. The EOMCCSD values of the dipole oscillator strengths characterizing the calculated electronic transitions, along with the available experimental [273] and CASSCF [262] data, can be found in Table 4.4.

As shown in Table 4.2, our  $\delta$ -CR-EOMCC(2,3) results for the low-lying valence excited states in azulene are in very good agreement with the available CASPT2(10,10) [262] and experimental [250, 251] data. In most cases, our  $\delta$ -CR-EOMCC(2,3) excitation energies corresponding to the low-lying valence states show a slightly better agreement with experiment than those obtained with CASPT2. The vertical excitation energies in Table 4.2 are accompanied by the reduced excitation level (REL) diagnostic introduced in Ref. [49], resulting

from EOMCCSD calculations, which is defined as (cf. Eq. (26) in Ref. [49])

$$\text{REL} = \frac{\sum_{n=0}^2 n \langle \phi | (R_{\mu,n})^\dagger R_{\mu,n} | \phi \rangle}{\sum_{n=0}^2 \langle \phi | (R_{\mu,n})^\dagger R_{\mu,n} | \phi \rangle} = \frac{\sum_{i,a} (r_a^i)^2 + 2 \sum_{i < j, a < b} (r_{ab}^{ij})^2}{(r_0)^2 + \sum_{i,a} (r_a^i)^2 + \sum_{i < j, a < b} (r_{ab}^{ij})^2}, \quad (4.27)$$

where, as explained in Section 4.1,  $r_0$ ,  $r_a^i$ , and  $r_{ab}^{ij}$  are the coefficients at the reference and the single and double excitation amplitudes, respectively, defining the one- and two-body components  $R_{\mu,1}$  and  $R_{\mu,2}$ , respectively, obtained in the EOMCCSD calculations, such that  $\text{REL} \approx 1$  implies a one-electron transition and  $\text{REL}$  close to 2 indicates a doubly excited state. As seen from their  $\text{REL}$  values in Table 4.2 being close to one, the four low-lying valence states of azulene,  $S_1$ – $S_4$ , are all predominantly one-electron transitions.

Our computations clearly show a doubly excited state with  $\text{REL} \approx 1.9$ , labeled in Tables 4.2–4.4 as state  $S_9$ , which satisfies all of the conditions set on it in Ref. [243]. Indeed, according to our  $\delta$ -CR-EOMCC(2,3) calculations, augmented with the EOMCCSD wave function information,  $S_9$  is a doubly excited  $^1A_1$  state dominated by the  $(\text{HOMO})^2 \rightarrow (\text{LUMO} + 1)^2$   $[(2a_2)^2 \rightarrow (3a_2)^2]$  and  $(\text{HOMO} - 1)^2 \rightarrow (\text{LUMO})^2$   $[(3b_2)^2 \rightarrow (4b_2)^2]$  transitions, and, as predicted from the experimental data summarized in Fig. 4.2, its vertical excitation energy of  $\sim 6.6$  eV lies between 5.19 and 6.81 eV (larger basis sets might lower the energy of  $S_9$  somewhat, but this would only strengthen our conclusions). Furthermore, as shown in Table 4.4, the oscillator strength of the  $S_2 \rightarrow S_9$  probe transition (0.004–0.005) is comparable to that characterizing the  $S_0 \rightarrow S_2$  pump transition (0.003), i.e., the calculated doubly excited state is accessible from  $S_2$ , as desired (with the  $S_2 \rightarrow S_9$  transition dipole oriented along the main rotational axis of the  $C_{2v}$ -symmetric azulene). All of this suggests that the doubly

excited state of azulene below the ionization threshold that supports the  $1 + 2'$  photoionization experiment [243] indeed exists. We need to emphasize that this high-lying state, in the middle of the Rydberg part of the spectrum, would not be found if we did not correct EOMCCSD, which is incapable of describing the energetics of doubly excited states in an accurate manner, for triple excitations to determine the electronic energy levels (using EOM-CCSDT would be even better, but such calculations for azulene are prohibitively expensive). On the other hand, EOMCCSD can provide useful information about dominant electronic configurations by examining the EOM excitation amplitudes, as in Table 4.3.

Table 4.1: Bond lengths (in Å) defining the calculated and experimental geometries of azulene in its ground electronic state  $S_0$  ( $X^1A_1$ ) (for the meaning of various distances, see Fig. 4.1).

| Bond     | MP2 <sup>a</sup> | MP2 <sup>b</sup> | CCSD <sup>b</sup> | CR-CC(2,3) <sup>a</sup> | CAS(10,10) <sup>c</sup> | B3LYP <sup>d</sup> | B3LYP <sup>e</sup> | X-ray <sup>f</sup> | ED <sup>g</sup> | $\mu$ -wave <sup>h</sup> |
|----------|------------------|------------------|-------------------|-------------------------|-------------------------|--------------------|--------------------|--------------------|-----------------|--------------------------|
|          | 6-31G(d)         | 6-311G(d,p)      | 6-31G(d,p)        | cc-pVDZ                 | 6-31G(d)                | 6-31G(d)           | cc-pVTZ            |                    |                 |                          |
| $r_1$    | 1.405            | 1.408            | 1.404             | 1.419                   | 1.405                   | 1.404              | 1.400              | 1.392              | 1.399           | 1.405                    |
| $r_2$    | 1.405            | 1.407            | 1.405             | 1.418                   | 1.404                   | 1.405              | 1.401              | 1.400              | 1.418           | 1.412                    |
| $r_3$    | 1.392            | 1.393            | 1.393             | 1.405                   | 1.392                   | 1.392              | 1.386              | 1.391              | 1.383           | 1.375                    |
| $r_4$    | 1.400            | 1.401            | 1.398             | 1.411                   | 1.400                   | 1.398              | 1.396              | 1.398              | 1.406           | 1.405                    |
| $r_5$    | 1.400            | 1.401            | 1.398             | 1.412                   | 1.400                   | 1.398              | 1.392              | 1.394              | 1.403           | 1.396                    |
| $r_c$    | 1.503            | 1.505            | 1.499             | 1.509                   | 1.497                   | 1.500              | 1.496              | 1.498              | 1.501           | 1.482                    |
| $r_{h1}$ | 1.087            |                  |                   | 1.096                   |                         | 1.085              | 1.080              |                    | 1.090           | 1.079                    |
| $r_{h2}$ | 1.086            |                  |                   | 1.095                   |                         | 1.084              | 1.079              |                    | 1.086           | 1.078                    |
| $r_{h3}$ | 1.092            |                  |                   | 1.101                   |                         | 1.090              | 1.086              |                    | 1.080           | 1.080                    |
| $r_{h4}$ | 1.089            |                  |                   | 1.099                   |                         | 1.088              | 1.083              |                    | 1.086           | 1.082                    |
| $r_{h5}$ | 1.090            |                  |                   | 1.100                   |                         | 1.089              | 1.084              |                    | 1.087           | 1.081                    |

<sup>a</sup> This work.

<sup>b</sup> Taken from Ref. [261].

<sup>c</sup> Taken from Ref. [262].

<sup>d</sup> Taken from Ref. [263].

<sup>e</sup> Taken from Ref. [264].

<sup>f</sup> X-ray diffraction data taken from Ref. [265].

<sup>g</sup> Gas-phase electron diffraction data taken from Ref. [266]. The vibrationally corrected C–H bond lengths taken from Ref. [267].

<sup>h</sup> Determined from the microwave spectra of azulene isotopomers taken from Ref. [268].

Table 4.2: Vertical excitation energies (in eV) and REL values for excited states of azulene found in this work.

| State         | Basis set | REL   | EOMCCSD | $\delta$ -CR-EOMCC(2,3) | CASSCF(10,10) <sup>a</sup> | CASPT2(10,10) <sup>a</sup> | Experiment <sup>b</sup> |
|---------------|-----------|-------|---------|-------------------------|----------------------------|----------------------------|-------------------------|
| $S_1(B_1)$    | 6-31G(d)  | 1.105 | 2.290   | 1.691                   | 1.97                       | 1.96                       | 1.77                    |
|               | cc-pVDZ   | 1.105 | 2.224   | 1.618                   |                            |                            |                         |
| $S_2(A_1)$    | 6-31G(d)  | 1.089 | 4.192   | 3.570                   | 4.45                       | 3.81                       | 3.56                    |
|               | cc-pVDZ   | 1.090 | 4.037   | 3.414                   |                            |                            |                         |
| $S_3(B_1)$    | 6-31G(d)  | 1.109 | 4.897   | 4.220                   | 4.62                       | 4.15                       | 4.23                    |
|               | cc-pVDZ   | 1.110 | 4.760   | 4.083                   |                            |                            |                         |
| $S_4(A_1)$    | 6-31G(d)  | 1.097 | 5.567   | 4.954                   | 5.50                       | 4.94                       | 4.40                    |
|               | cc-pVDZ   | 1.097 | 5.387   | 4.770                   |                            |                            |                         |
| $S_5(B_1)$    | 6-31G(d)  | 1.107 | 6.424   | 5.792                   |                            |                            | 4.72                    |
|               | cc-pVDZ   | 1.107 | 6.237   | 5.607                   |                            |                            |                         |
| $S_6(A_1)$    | 6-31G(d)  | 1.146 | 6.671   | 5.824                   |                            |                            | 5.19                    |
|               | cc-pVDZ   | 1.143 | 6.509   | 5.674                   |                            |                            |                         |
| $S_7(B_2)$    | 6-31G(d)  | 1.093 | 6.543   | 6.010                   |                            |                            | 6.15                    |
|               | cc-pVDZ   | 1.095 | 6.421   | 5.888                   |                            |                            |                         |
| $S_8(A_1)$    | 6-31G(d)  | 1.140 | 7.378   | 6.473                   |                            |                            | 6.33                    |
|               | cc-pVDZ   | 1.138 | 7.187   | 6.298                   |                            |                            |                         |
| $S_9(A_1)$    | 6-31G(d)  | 1.879 | 10.708  | 6.787                   |                            |                            |                         |
|               | cc-pVDZ   | 1.893 | 10.652  | 6.578                   |                            |                            |                         |
| $S_{10}(A_2)$ | 6-31G(d)  | 1.092 | 7.348   | 6.810                   |                            |                            |                         |
|               | cc-pVDZ   | 1.094 | 7.186   | 6.648                   |                            |                            |                         |

<sup>a</sup>From Ref. [262].

<sup>b</sup>From Table 1 in Ref. [243]. Experiment points to the existence of the Rydberg states at 4.72 ( $R_A$ ), 5.19 ( $R_B$ ), 5.64 ( $R_C$ ), 5.92 ( $R_D$ ), 6.15 ( $R_E$ ), and 6.33 ( $R_F$ ) eV. Based on the orbital character of the leading excitation amplitudes, states  $R_A$ ,  $R_B$ ,  $R_E$ , and  $R_F$  from Table 1 in Ref. [243] can be assigned to the EOMCC roots  $S_5$ ,  $S_6$ ,  $S_7$ , and  $S_8$ , respectively. States  $R_C$  and  $R_D$  which have  $3d$  and  $3d/4s$  character, respectively [243], and which might, therefore, be sensitive to the molecular geometry [259] that was not optimized for excited states, could not be found among the calculated EOMCC roots, likely due to the inadequacy of the basis sets employed.

Table 4.3: The absolute values of the reference ( $r_{\mu,0}$ ) and leading singly excited ( $r_a^i$ ) and doubly excited ( $r_{ab}^{ij}$ ) amplitudes, along with the corresponding orbital information, defining the EOMCCSD wave functions of various excited states of azulene obtained using the 6-31G(d) and cc-pVDZ basis sets.

| State         | Basis set | $r_{\mu,0}$ | Leading $r_a^i$ and $r_{ab}^{ij}$ amplitudes <sup>a</sup> |                                      |                                   |                                         |                                         |                                                                     |
|---------------|-----------|-------------|-----------------------------------------------------------|--------------------------------------|-----------------------------------|-----------------------------------------|-----------------------------------------|---------------------------------------------------------------------|
| $S_1(B_1)$    | 6-31G(d)  | 0.00        | $2a_2 \rightarrow 4b_2(0.64)$                             | $3b_2 \rightarrow 3a_2(0.18)$        |                                   |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.00        | $2a_2 \rightarrow 4b_2(0.64)$                             | $3b_2 \rightarrow 3a_2(0.17)$        |                                   |                                         |                                         |                                                                     |
| $S_2(A_1)$    | 6-31G(d)  | 0.02        | $3b_2 \rightarrow 4b_2(0.45)$                             | $2a_2 \rightarrow 3a_2(0.49)$        |                                   |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.03        | $3b_2 \rightarrow 4b_2(0.45)$                             | $2a_2 \rightarrow 3a_2(0.49)$        |                                   |                                         |                                         |                                                                     |
| $S_3(B_1)$    | 6-31G(d)  | 0.00        | $1a_2 \rightarrow 4b_2(0.30)$                             | $2a_2 \rightarrow 4b_2(0.15)$        | $3b_2 \rightarrow 3a_2(0.56)$     |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.00        | $1a_2 \rightarrow 4b_2(0.30)$                             | $2a_2 \rightarrow 4b_2(0.14)$        | $3b_2 \rightarrow 3a_2(0.56)$     |                                         |                                         |                                                                     |
| $S_4(A_1)$    | 6-31G(d)  | 0.04        | $3b_2 \rightarrow 4b_2(0.47)$                             | $1a_2 \rightarrow 3a_2(0.13)$        | $2a_2 \rightarrow 3a_2(0.43)$     |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.04        | $3b_2 \rightarrow 4b_2(0.47)$                             | $1a_2 \rightarrow 3a_2(0.13)$        | $2a_2 \rightarrow 3a_2(0.43)$     |                                         |                                         |                                                                     |
| $S_5(B_1)$    | 6-31G(d)  | 0.00        | $1a_2 \rightarrow 4b_2(0.53)$                             | $2b_2 \rightarrow 3a_2(0.18)$        | $3b_2 \rightarrow 3a_2(0.24)$     | $2a_2 \rightarrow 5b_2(0.20)$           |                                         |                                                                     |
|               | cc-pVDZ   | 0.00        | $1a_2 \rightarrow 4b_2(0.53)$                             | $2b_2 \rightarrow 3a_2(0.18)$        | $3b_2 \rightarrow 3a_2(0.24)$     | $2a_2 \rightarrow 5b_2(0.21)$           |                                         |                                                                     |
| $S_6(A_1)$    | 6-31G(d)  | 0.03        | $2b_2 \rightarrow 4b_2(0.51)$                             | $1a_2 \rightarrow 3a_2(0.36)$        | $3b_2 \rightarrow 5b_2(0.11)$     | $3b_2^2 \rightarrow 4b_2^2(0.12)$       | $2a_2^2 \rightarrow 3a_2^2(0.12)$       |                                                                     |
|               | cc-pVDZ   | 0.03        | $2b_2 \rightarrow 4b_2(0.51)$                             | $1a_2 \rightarrow 3a_2(0.37)$        | $3b_2 \rightarrow 5b_2(0.11)$     | $3b_2^2 \rightarrow 4b_2^2(0.12)$       | $2a_2^2 \rightarrow 3a_2^2(0.12)$       |                                                                     |
| $S_7(B_2)$    | 6-31G(d)  | 0.00        | $17a_1 \rightarrow 4b_2(0.66)$                            |                                      |                                   |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.00        | $17a_1 \rightarrow 4b_2(0.66)$                            |                                      |                                   |                                         |                                         |                                                                     |
| $S_8(A_1)$    | 6-31G(d)  | 0.04        | $2b_2 \rightarrow 4b_2(0.34)$                             | $1a_2 \rightarrow 3a_2(0.49)$        | $2a_2 \rightarrow 3a_2(0.12)$     | $3b_2 \rightarrow 5b_2(0.11)$           | $2a_2 \rightarrow 4a_2(0.14)$           | $2a_2^2 \rightarrow 3a_2^2(0.13)$                                   |
|               | cc-pVDZ   | 0.04        | $2b_2 \rightarrow 4b_2(0.34)$                             | $1a_2 \rightarrow 3a_2(0.49)$        | $2a_2 \rightarrow 3a_2(0.12)$     | $3b_2 \rightarrow 5b_2(0.11)$           | $2a_2 \rightarrow 4a_2(0.14)$           | $2a_2^2 \rightarrow 3a_2^2(0.12)$                                   |
| $S_9(A_1)$    | 6-31G(d)  | 0.15        | $3b_2 \rightarrow 6b_2(0.10)$                             | $1a_2 2a_2 \rightarrow 3a_2^2(0.12)$ | $3b_2^2 \rightarrow 4b_2^2(0.41)$ | $3b_2 2a_2 \rightarrow 4b_2 3a_2(0.23)$ | $3b_2 2a_2 \rightarrow 3a_2 4b_2(0.28)$ | $2a_2^2 \rightarrow 4b_2^2(0.11)$ $2a_2^2 \rightarrow 3a_2^2(0.42)$ |
|               | cc-pVDZ   | 0.15        |                                                           | $1a_2 2a_2 \rightarrow 3a_2^2(0.12)$ | $3b_2^2 \rightarrow 4b_2^2(0.41)$ | $3b_2 2a_2 \rightarrow 4b_2 3a_2(0.23)$ | $3b_2 2a_2 \rightarrow 3a_2 4b_2(0.28)$ | $2a_2^2 \rightarrow 4b_2^2(0.11)$ $2a_2^2 \rightarrow 3a_2^2(0.42)$ |
| $S_{10}(A_2)$ | 6-31G(d)  | 0.00        | $12b_1 \rightarrow 4b_2(0.65)$                            |                                      |                                   |                                         |                                         |                                                                     |
|               | cc-pVDZ   | 0.00        | $12b_1 \rightarrow 4b_2(0.65)$                            |                                      |                                   |                                         |                                         |                                                                     |

<sup>a</sup>Only the amplitudes whose absolute values are greater than 0.1 are shown.

Table 4.4: Oscillator strengths for the various transitions in azulene obtained from the EOMCCSD calculations using the 6-31G(d) and cc-pVDZ basis sets and the available CASSCF(10,10)/6-31G(d) [262] and experimental [273] results.

| State         | Basis         | $S_0(A_1)$ | $S_1(B_1)$ | $S_2(A_1)$ | $S_3(B_1)$ | $S_4(A_1)$ | $S_5(B_1)$ | $S_6(A_1)$ | $S_7(B_2)$ | $S_8(A_1)$ | $S_9(A_1)$ |
|---------------|---------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| $S_1(B_1)$    | 6-31G(d)      | 0.007      |            |            |            |            |            |            |            |            |            |
|               | cc-pVDZ       | 0.007      |            |            |            |            |            |            |            |            |            |
|               | CASSCF(10.10) | 0.003      |            |            |            |            |            |            |            |            |            |
|               | Experiment    | 0.009      |            |            |            |            |            |            |            |            |            |
| $S_2(A_1)$    | 6-31G(d)      | 0.003      | 0.000      |            |            |            |            |            |            |            |            |
|               | cc-pVDZ       | 0.003      | 0.000      |            |            |            |            |            |            |            |            |
|               | CASSCF(10.10) | 0.002      |            |            |            |            |            |            |            |            |            |
|               | Experiment    | 0.08       |            |            |            |            |            |            |            |            |            |
| $S_3(B_1)$    | 6-31G(d)      | 0.064      | 0.022      | 0.001      |            |            |            |            |            |            |            |
|               | cc-pVDZ       | 0.062      | 0.022      | 0.001      |            |            |            |            |            |            |            |
|               | CASSCF(10.10) | 0.014      |            |            |            |            |            |            |            |            |            |
| $S_4(A_1)$    | 6-31G(d)      | 1.309      | 0.008      | 0.001      | 0.000      |            |            |            |            |            |            |
|               | cc-pVDZ       | 1.307      | 0.008      | 0.001      | 0.000      |            |            |            |            |            |            |
|               | CASSCF(10.10) | 0.067      |            |            |            |            |            |            |            |            |            |
|               | Experiment    | 1.10       |            |            |            |            |            |            |            |            |            |
| $S_5(B_1)$    | 6-31G(d)      | 0.235      | 0.038      | 0.020      | 0.007      | 0.000      |            |            |            |            |            |
|               | cc-pVDZ       | 0.226      | 0.038      | 0.020      | 0.006      | 0.000      |            |            |            |            |            |
|               | Experiment    | 0.38       |            |            |            |            |            |            |            |            |            |
| $S_6(A_1)$    | 6-31G(d)      | 0.019      | 0.036      | 0.119      | 0.000      | 0.015      | 0.000      |            |            |            |            |
|               | cc-pVDZ       | 0.018      | 0.036      | 0.125      | 0.000      | 0.014      | 0.000      |            |            |            |            |
| $S_7(B_2)$    | 6-31G(d)      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      |            |            |            |
|               | cc-pVDZ       | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      |            |            |            |
| $S_8(A_1)$    | 6-31G(d)      | 0.228      | 0.023      | 0.000      | 0.017      | 0.039      | 0.003      | 0.000      | 0.000      |            |            |
|               | cc-pVDZ       | 0.229      | 0.025      | 0.000      | 0.017      | 0.041      | 0.003      | 0.000      | 0.000      |            |            |
|               | Experiment    | 0.65       |            |            |            |            |            |            |            |            |            |
| $S_9(A_1)$    | 6-31G(d)      | 0.001      | 0.043      | 0.005      | 0.052      | 0.190      | 0.014      | 0.002      | 0.000      | 0.034      |            |
|               | cc-pVDZ       | 0.000      | 0.043      | 0.004      | 0.038      | 0.212      | 0.015      | 0.003      | 0.000      | 0.023      |            |
| $S_{10}(A_2)$ | 6-31G(d)      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      |
|               | cc-pVDZ       | 0.000      | 0.001      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      | 0.000      |



Our discovery of the doubly excited state of azulene below the ionization threshold, which can mediate multiphoton ionization may have substantial implications for the rapidly developing field of Rydberg fingerprint spectroscopy (RFS) [259,274–278], since similar states may be present and be equally easily accessible in other molecular systems. The doubly excited states are dark, when one tries to populate them from the ground state, but, as shown in Ref. [243] and this study [142], they can be populated via transitions from low-lying bright states, allowing one to prepare molecules in their respective Rydberg levels via fast internal conversion, prior to ionization, which is the basis of RFS. One of the advantages of RFS, related to the fact that ionization from a Rydberg state is essentially purely electronic and uncoupled from vibrational transitions, is its remarkable sensitivity to structural changes, enabling one to investigate the time-resolved conformational dynamics in vibrationally hot floppy molecules, a very difficult task to achieve with other experimental techniques (cf., e.g., Refs. [279,280]). Using doubly excited states below the ionization threshold, such as that found here, as a way to prepare molecules in their respective Rydberg levels may open up new avenues in this important direction.

## 4.3 Aerobic oxidation of methanol to formic acid on $\text{Au}_8^-$ : Benchmark analysis based on completely renormalized coupled-cluster and density functional theory calculations

### 4.3.1 Background information and scope of the work

The pioneering work of Haruta and coworkers [200] and Hutchings and coworkers [199], investigating the low-temperature oxidation of CO and the hydrochlorination of ethyne to vinyl chloride, respectively, has led to an increased interest in the catalytic properties of small gold nanoparticles from both theory and experiment [201–207, 281–284]. Many useful reactions have been developed using various types of gold catalysts, such as anchored gold nanoparticles on metal oxides and dispersions of colloidal gold in liquid media. Among the various reactions, aerobic oxidations have received much attention due to their high yields even when carried out under ambient conditions, yielding the desired products in a highly selective manner.

As already mentioned in Section 3.3.1, the chemistry of gold has been extensively studied by various groups, using a variety of computational approaches (cf. Refs. [204–207] for representative reviews). In particular, a great deal of attention has been dedicated to the electronic structure, geometries, and catalytic activity of gold clusters, most of which have been primarily studied using methods based on DFT. As also pointed out in Section 3.3.1, the application of reliable *ab initio* wave function approaches, in particular, the size-extensive

methods based on CC theory, such as CCSD(T), CCSD(2)<sub>T</sub>, or CR-CC(2,3), where energies obtained with the basic CCSD approach are corrected for the dominant effects due to connected triple excitations to obtain the desired high accuracies, have been very limited to date, due in large part to the high computational costs for systems involving polyatomic gold particles. The low energy structures of Au<sub>6</sub> and Au<sub>8</sub> were studied by Olson *et al.* [214] at the CCSD(T) level, using the CCSD(T) routines in GAMESS developed by the Piecuch group, comparing their CCSD(T) results with those obtained using the DFT and MP2 approaches. Diefenbach and Kim [215], Han [216], Olson and Gordon [217], and most recently, our group [183] examined the effect of the basis set, core correlations, and geometry relaxation in Au<sub>8</sub> using CCSD(T) (see Section 5.1.3). Varganov *et al.* investigated the adsorption of O<sub>2</sub> (Ref. [211]) and H<sub>2</sub> (Ref. [213]) and their activation on small Au<sub>n</sub> and Au<sub>n</sub><sup>-</sup> ( $n = 2, 3$ ) clusters employing the DFT, MP2, and CCSD(T) methods. However, very little is known about the performance of the high-level methods based on CC theory in applications involving more complex reactions catalyzed by gold particles. Our work reported in this section and Ref. [181] address this concern.

In the case of larger gold clusters, the various DFT results are generally consistent with experiment [285], but the overall situation, particularly when reactive processes catalyzed by gold clusters are considered, is far from ideal. The underlying reaction mechanisms continue to be debated [286–289], since neither DFT, where the results often depend (in many cases strongly) on the applied functional, nor experiments, which give useful but only partial insights, can provide definitive information. For this reason, calculations comparing DFT methods, which are applicable to large molecular systems, with the robust wave function

approaches that can provide accurate benchmark data are in growing demand. In this thesis project, reported in Ref. [181], we undertook the study of the complex problem of aerobic oxidation of methanol to formic acid catalyzed by  $\text{Au}_8^-$  at the high *ab initio* CCSD(2)<sub>T</sub> and CR-CC(2,3) levels, for the first time ever. As already alluded to above (cf., e.g., the introduction in Chapter 1) these methods improve the performance of the popular CCSD(T) approach in multireference situations, where CCSD(T) fails, including bond breaking [51,68,77,93–95,99,290,291] (cf. also, Section 5.2), chemical reaction pathways involving biradicals [98,105,292], open-shell species [95,291], and larger transition-metal containing species [99,105]. By having access to the high-level CCSD(2)<sub>T</sub> and CR-CC(2,3) data, we can assess the performance of the representative DFT methods.

In general, in aerobic oxidation reactions on gold nanoparticles, the activation of molecular oxygen is a critical step [293,294] with the electronic charge transfer from the gold particles to the adsorbed oxygen molecules playing a crucial role [295]. However, other factors may contribute too. For example, Okumura and coworkers examined the significance of the heterojunction effect between the polymerlike poly(vinyl-2-pyrrolidone) (PVP) system used in the experimental work on gold-catalyzed aerobic oxidation reactions and Au clusters [287,296]. Lyalin and Taketsuga reported that cooperative effects might be important as well [297,298]. The electronic structure aspects of aerobic oxidation on gold particles and the usefulness of the CC and DFT methods in examining some of these aspects are focuses of this thesis project.

Recently, Ehara and coworkers investigated the aerobic oxidation of methanol to formic acid on gold particles supported by PVP (Au:PVP) using DFT [299,300] exploiting the M06

functional [301]. Two reaction pathways with relatively low activation barriers have been suggested. To design suitable theoretical models of the catalytic site, Ehara and coworkers adopted the negatively charged  $\text{Au}_8^-$  (Ref. [300]) and  $\text{Au}_{20}^-$  (Ref. [299]) clusters, taking into account the electron-donating nature of the PVP medium, the dependence of the electron affinity (EA) of the  $\text{Au}_n$  clusters on  $n$ , which exhibits minima at  $n = 8$  and  $20$  (Ref. [221]), and the fact that clusters with low EAs activate the oxidative reactions most effectively [302]. The smooth exothermic reaction pathways were obtained and the stationary points defining these pathways were determined via full geometry optimizations (including all gold nanoparticle and reactant degrees of freedom) using the M06 functional. They showed that the reactivity of the smaller planar  $\text{Au}_8^-$  cluster is not as high as in the case of the larger non-planar  $\text{Au}_{20}^-$  species, in agreement with the experimentally observed enhancement of reactivity when the gold particles in the Au:PVP system catalyzing the aerobic oxidation of methanol to formic acid in the  $\sim 1$  nm range [202] and consistent with surface roughening arguments mentioned later in Section 5.1.3.

As one can see, the aerobic oxidation of methanol to formic acid on the Au:PVP system has already been theoretically analyzed, suggesting the plausible reaction mechanisms, but this is all based on uncertain DFT data obtained with a single functional and rather small basis sets. In order to extend such studies to other aerobic oxidation processes catalyzed by gold and other metallic particles, it is essential to verify whether the previously reported M06 calculations [299,300] are trustworthy. There are numerous aerobic oxidation reactions using monometallic, bimetallic [303], or even trimetallic [304] nanoparticles whose mechanisms await understanding. For example, recently the unique coupling reaction for

Cl–Cl bond activation with Au/Pd alloy nanoparticles was reported, where the reaction on pure Au or Pd nanoparticles does not proceed at all, but the Au/Pd alloy system displays catalytic activity [305]. Clearly, findings of this kind may depend on the employed DFT functional. It is, therefore, important to determine how accurate various DFT approaches are in similar applications by testing them against robust *ab initio* theories than can offer a reliable description of chemical reaction pathways, which is the focus of this part of the dissertation.

In this thesis project, we examine the energetics of methanol oxidation to formic acid on  $\text{Au}_8^-$  using, as the highest theory level for benchmarking DFT, the CR-CC(2,3) approach, which provides an accurate description of the dynamical and nondynamical correlation effects relevant to single bond breaking, including singlet [51, 68, 93, 94, 99, 290] (cf. Section 5.2), and nonsinglet [95, 291] potential energy surfaces, reaction pathways involving biradicals [98, 105, 292], and transition-metal-containing species relevant to catalytic activity [99, 105], with the ease of a single-reference black-box calculation. We also use the CCSD(2)<sub>T</sub> approach, which, as explained in Refs. [93–95] and Section 4.1, is one of the variants of CR-CC(2,3) and which eliminates, in analogy to CR-CC(2,3), the failures of CCSD(T) in bond-breaking situations. We focus on the most probable reaction mechanism, labeled in the previous studies [299, 300] as pathway I, in which formic acid is produced by oxidation of the methoxy species ( $\text{CH}_3\text{O}^-$ ) on  $\text{Au}_8^-$ . The resulting CR-CC(2,3) pathway is used to benchmark the representative DFT approaches, which in addition to the previously employed [299, 300] M06 hybrid functional include its pure generalized gradient approximation (GGA)-type M06-L analog [301], two other pure GGAs, namely, BP86 [306, 307] and

TPSS [308], the dispersion-corrected counterpart of BP86 represented by the B97-D functional [309], and the popular B3LYP hybrid scheme [310]. Additionally, in this dissertation, we include the unpublished results obtained using the dispersion corrected hybrid functional  $\omega$ -B97X-D [311] for comparison with our results published in Ref. [181].

### 4.3.2 Molecular model

We want to be able to describe the mechanism of the oxidation of methanol to formic acid on the gold clusters, which are the most essential part of the real Au:PVP system, using high-level CC methods, such as CR-CC(2,3), and a few representative DFT approaches, and thus, to come up with the appropriate molecular model amenable to quantum chemistry computation, we need to take into consideration the known properties of small gold nanoparticles and other experimentally determined facts concerning the Au:PVP-catalyzed aerobic oxidation of alcohols. Due to the electron-donating property of PVP in the Au:PVP systems [287], we will use negatively charged ( $\text{Au}_n^-$ ) rather than neutral ( $\text{Au}_n$ ) species. Of the various gold clusters, at least those with up to 58 Au atoms, the best candidates for activating molecular oxygen adsorbed on the Au:PVP system are the negatively charged  $\text{Au}_8^-$  and  $\text{Au}_{20}^-$  species, whose corresponding neutral clusters show the strongest drop in their corresponding EA values [221]. This is in agreement with the study by Salisbury *et al.* [302], where they indicated that the  $\text{Au}_8^-$  and  $\text{Au}_{20}^-$  particles are the most effective electron donors in the process of activating  $\text{O}_2$  among the nanoclusters in the  $\sim 1$  nm range, which yields the optimum aerobic oxidation of alcohols [312]. It would obviously be useful to examine both the  $\text{Au}_8^-$  and  $\text{Au}_{20}^-$  species in the aerobic oxidation reaction considered in this dissertation, but the

CR-CC(2,3) calculations involving  $\text{Au}_{20}^-$  are prohibitively expensive for us at this time if we do not use the local formulation of CR-CC(2,3). Thus, our investigation of the oxidation of methanol to formic acid on gold nanoparticles carried out as part of this thesis research with high-level CC methods, such as CR-CC(2,3), we will focus on the smaller model involving the  $\text{Au}_8^-$  cluster.

In the previous DFT/M06 studies of methanol oxidation to formic acid on  $\text{Au}_8^-$  (Ref. [300]), two reaction pathways were considered. The first of the two, called pathway I, which we believe is the most likely one and which is schematically shown in Fig. 4.4, all of the oxidation steps on the  $\text{Au}_8^-$  cluster occur sequentially. In this case, the reaction begins by coadsorption of the oxygen molecule and the methoxy species, which is generated from methanol in the basic environment used in the experiments, on the  $\text{Au}_8^-$  particle to form structure A1. The hydrogen of the methoxy species is subsequently transferred to the molecular oxygen via transition state TS\_A1 to form structure B1 containing formaldehyde. In the following step, intrasystem OH transfer via transition state TS\_B1 occurs, forming the hemiacetal intermediate C1, and in the final stages of the reaction, intrasystem hydrogen transfer in C1 via the transition state TS\_C1 leads to the release of the products, namely, formic acid and the  $(\text{Au}_8\text{OH})^{2-}$  species E1, via the intermediate complex D1. The  $(\text{Au}_8\text{OH})^{2-}$  adduct decomposes into its corresponding pieces,  $\text{Au}_8^-$ , which returns to the catalytic cycle, and  $\text{OH}^-$ , which returns to the basic solution. The second pathway considered in Refs. [299,300], which we believe is less likely, involves detachment of the formaldehyde after the B1 intermediate is formed, followed by generation of the hemiacetal on the Au cluster. In this dissertation, as mentioned above, we focus on the more probable pathway I shown in Fig. 4.4.



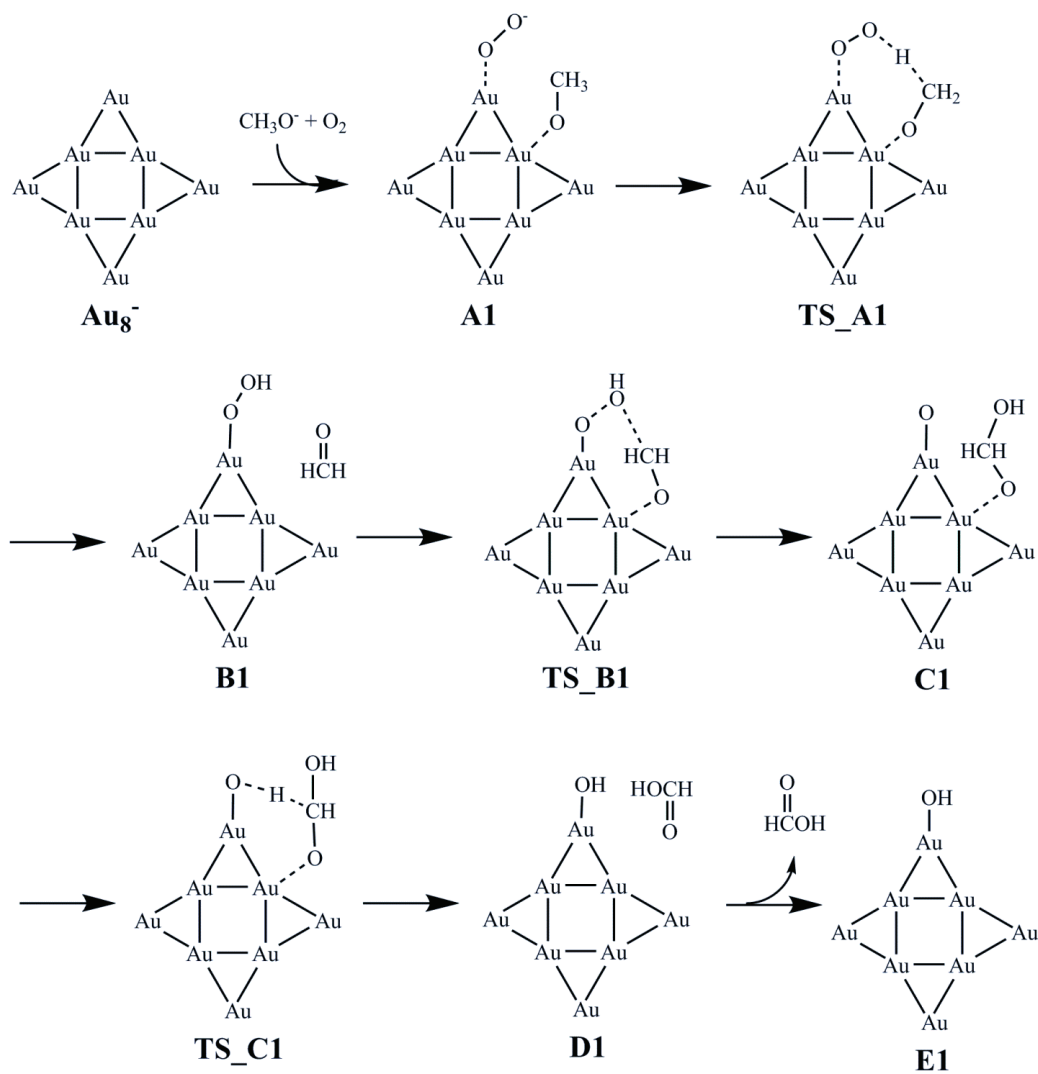


Figure 4.4: Schematic representation of the reaction pathway I for the methanol oxidation to formic acid on the  $\text{Au}_8^-$  cluster proposed in Ref. [300].

### 4.3.3 Electronic structure methods

In order to assess the performance of the selected DFT approaches with reliable *ab initio* benchmark data, we carried out single-point CR-CC(2,3) calculations using the geometries of the isolated reactants ( $\text{CH}_3\text{O}^-$ ,  $\text{O}_2$ , and  $\text{Au}_8^-$ ), products ( $\text{HCOOH}$  and  $(\text{Au}_8\text{OH})^{2-}$ ), and the remaining stationary points A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, and D1 defining pathway I obtained in the earlier DFT/M06 study [300]. As explained earlier, the CR-CC(2,3) approach represents a rigorously size-extensive method that can provide an accurate description of single-bond breaking, biradicals, and reaction pathways involving transition-metal systems with an effort similar to conventional CCSD(T) calculations, while eliminating the failures of CCSD(T) when bonds are stretched or broken (cf. Section 5.2 below). In this thesis project, we focused on the most complete and typically most accurate CR-CC(2,3),D variant, which uses the Epstein-Nesbet-like form of the perturbative denominator defining the triples correction of CR-CC(2,3), Eq. (4.22), and the CR-CC(2,3),A approximation, which uses the Møller-Plesset form of this denominator, Eq. (4.23), and which is equivalent to the CCSD(2)<sub>T</sub> approach when the canonical Hartree-Fock (HF) orbitals are used. Since both of these CR-CC(2,3) approaches rely on the singly and doubly excited clusters obtained in the CCSD calculations, the CCSD results are reported as well. The CCSD energies obtained with two different basis sets, including the smaller basis set of polarized double- $\zeta$  quality employed in the earlier M06 work [300] and its larger triple- $\zeta$  counterpart, combined with the CR-CC(2,3) calculations using the basis employed in Ref. [300], are also needed to extrapolate the larger basis set CR-CC(2,3) data, as described below. The most complete variant of CR-CC(2,3), namely, CR-CC(2,3),D, is treated as the method providing reference

energies against which the other CC and DFT approaches examined in this dissertation are benchmarked.

In the earlier DFT/M06 work [299,300], it was shown that all of the stationary points defining pathway I are doublet states. Therefore, all of the CC calculations reported in this section of this dissertation were performed using the doublet ROHF references. We confirmed, through suitable HF stability analysis, that the doublet ROHF solutions used in this work are, in fact, the lowest-energy HF solutions maintaining the spin symmetry. It is well-known that there are several methods of obtaining canonical ROHF orbitals that differ in the way the diagonal doubly occupied, singly occupied, and unoccupied blocks of the Fock matrix are constructed (see, e.g., Refs. [313,314]). While iterative CC methods, such as CCSD, are invariant with respect to the canonicalization of the ROHF orbitals, they are not invariant with respect to orbital rotations among the occupied orbitals and unoccupied orbitals when one freezes core orbitals, which is the case in the calculations carried out in this study. Furthermore, the ROHF-based CR-CC(2,3) triples corrections, as implemented in the GAMESS package by the Piecuch group, may show a small dependence on the way the ROHF orbitals are generated [95]. We, thus, considered two ways to generate the ROHF references for use in our CC calculations, namely, Roothaan’s scheme [313], which is the default in GAMESS, and the equally popular Guest-Saunders approach [314], though, based on the results of Ref. [95] we did not anticipate this to be a major issue here.

As mentioned above, the energetics resulting from our CR-CC(2,3),D calculations characterizing pathway I were used to benchmark a few representative DFT approaches, including the previously employed [299,300] M06 hybrid functional, the pure GGA-type analog of M06

abbreviated as M06-L [301], the BP86 [306,307] and TPSS [308] pure GGAs, the dispersion-corrected B97-D scheme [309], the popular B3LYP hybrid approach [310], and the long-range corrected hybrid  $\omega$ -B97X-D [311] functional. In analogy to the CC calculations, all of the DFT computations reported in this portion of this dissertation were single-point calculations at the M06 geometries obtained in the previous study [300] and performed with GAMESS. As usual, the DFT results reported in this work were obtained by the unrestricted Kohn-Sham computations.

In order to demonstrate that the main conclusions of this study, carried out as part of this thesis research, are not affected by the choice of the one-electron basis set, two different basis sets, designated as BS1 and BS2, were employed in our calculations. In the most complete set of computations, including all of the above DFT functionals and the CCSD, CR-CC(2,3),A  $\equiv$  CCSD(2)<sub>T</sub>, and CR-CC(2,3),D methods, we adopted the same smaller basis set used in the previous work [300], abbreviated as BS1, which combines the LANL2DZ [315] description of the gold atoms with the 6-31++G(d,p) basis set [316–318] for the remaining hydrogen, carbon, and oxygen atoms. We recall that the LANL2DZ description of the gold atom represents the inner shells corresponding to the (1s2s2p3s3p3d4s4p4d4f) core by the scalar relativistic ECP, while the remaining 19 electrons of the outer 5s, 5p, 5d, and 6s shells, which in the SCF (DFT and ROHF here) calculations are treated explicitly, are described by the double- $\zeta$  ([3s3p2d]) basis. As is usually done and to keep the computational costs at a manageable level, in the post-SCF stages of the CC calculations the 1s orbitals of the carbon and oxygen atoms and the 5s and 5p orbitals of the gold atoms outside the relativistic core were frozen. Thus, in our final CC calculations, we correlated the 58  $\alpha$  and 57  $\beta$  electrons

corresponding to the valence 5d and 6s shells of the Au atoms, the 2s and 2p shells of the C and O atoms, the 1s shells of the H atoms, and two extra electrons due to the negative charges of the  $\text{Au}_8^-$  and methoxy species.

Because the post-SCF wave function approaches, such as those based on CC theory, do not converge with the one-electron basis set as rapidly as the Kohn-Sham DFT (and other SCF-type) methods and since the LANL2DZ basis set may be somewhat too small to guarantee an accurate description of smaller gold particles (see, e.g., Refs. [183, 209, 215–217]), in addition to BS1 we also employed the considerably larger BS2 basis, in which the LANL2TZ(f) ECP and basis set developed for gold [319] are combined with the 6-311++G(d,p) basis set for the hydrogen, carbon, and oxygen atoms [318, 320]. As explained in Ref. [319], in the LANL2TZ(f) description of the gold atom, the inner shells corresponding to the (1s2s2p3s3p3d4s4p4d4f) core continue to be treated by the relativistic ECP (the same one as used in the LANL2DZ case), but the remaining 19 electrons of the outer 5s, 5p, 5d, and 6s shells, which are treated explicitly, are described by the triple- $\zeta$ -quality [5s5p3d] basis augmented by a single primitive f function as determined in Ref. [321].

When the smaller BS1 basis set was employed in the calculations, the  $\text{CH}_3\text{O}^- + \text{O}_2 + \text{Au}_8^-$  reactive system investigated as part of this work became (after removal of the relativistic cores of the gold atoms and the use of spherical components of the d functions) a 187-electron problem described by 266 basis functions, which, after freezing the 1s orbitals of the carbon and oxygen atoms and the 5s and 5p orbitals of the gold atoms, was manageable at all CC levels considered here, including the highest-level CR-CC(2,3),A and CR-CC(2,3),D approaches. Unfortunately, when describing the  $\text{CH}_3\text{O}^- + \text{O}_2 + \text{Au}_8^-$  system at the CR-

CC(2,3),A and CR-CC(2,3),D levels of theory, in which 115 electrons must be explicitly correlated, using the considerably larger BS2 basis set consisting after elimination of the relativistic core of each gold atom and the use of spherical d and f orbitals) of 445 functions, turned out to be prohibitively expensive for us at this time. Thus, in order to estimate the desired CR-CC(2,3),X/BS2 ( $X = \text{A, D}$ ) energies, we combined the results of the larger-basis-set ROHF-based CCSD/BS2 calculations, which we could perform in parallel by exploiting the NWChem package [322] (utilizing 28 cores for every stationary point along pathway I), with the corresponding smaller-basis-set CR-CC(2,3),X/BS1 ( $X = \text{A, D}$ ) information, generated with GAMESS, as follows:

$$\text{CR-CC}(2, 3), X/\text{BS2} = \text{CCSD}/\text{BS2} + [\text{CR-CC}(2, 3), X/\text{BS1} - \text{CCSD}/\text{BS1}]. \quad (4.28)$$

In other words, in estimating the highest-level CR-CC(2,3),A/BS2 and CR-CC(2,3),D/BS2 energetics, we corrected the CCSD/BS2 energies by the contributions due to connected triple excitations resulting from the corresponding CR-CC(2,3) calculations in the smaller BS1 basis set. In particular, the highest-level CR-CC(2,3),D/BS2 energies reported in this thesis work, which were used to benchmark various DFT approaches, were extrapolated using Eq. (4.28).

Since larger-basis-set CCSD calculations for catalytic systems with multiple transition-metal atoms may not always be possible, we also examined the possibility of estimating the larger-basis-set CCSD/BS2, CR-CC(2,3),A/BS2, and CR-CC(2,3),D/BS2 energies using information about the effect of going from the BS1 basis to its larger BS2 counterpart extracted from relatively inexpensive unrestricted MP2 (UMP2) calculations employing these

two basis sets (the results published in Ref. [181]) and the still inexpensive unrestricted third-order Møller-Plesset (UMP3) computations (the unpublished data). In both of these cases, the desired CC/BS2 energies were extrapolated using the formula

$$\text{CC/BS2} = \text{CC/BS1} + [\text{UMP}_n/\text{BS2} - \text{UMP}_n/\text{BS1}], \quad (4.29)$$

where CC stands for CCSD, CR-CC(2,3),A, or CR-CC(2,3),D and  $\text{UMP}_n$  stands for UMP2, when  $n = 2$ , and UMP3, when  $n = 3$ . A similar approach in the context of CR-CC(2,3),D calculations for various bond-breaking reactions proceeding on nonsinglet potential energy surfaces relevant to the gas-phase chemistry in silicon carbide chemical vapor deposition has previously been tested, obtaining in each case as accurate representation of the large-basis-set CR-CC(2,3),D data [291], but the effectiveness of Eq. (4.29) had not been examined before the work carried out as a part of this dissertation in applications of CR-CC(2,3) relevant to catalysis. Since we have access to the true CCSD/BS2 information, we also have an opportunity to examine whether the extrapolated (using Eq. (4.29)) and calculated CCSD/BS2 energies characterizing pathway I shown in Fig. 4.4 agree. If they do agree, it may reasonably be assumed that a similar approach applies to the CR-CC(2,3) energies, that is that we may use Eq. (4.29) to extrapolate the larger-basis-set CR-CC(2,3)/BS2 information from the smaller-basis-set CR-CC(2,3)/BS1 data and the UMP2/UMP3 energies obtained with the BS1 and BS2 basis sets, which are much easier to determine than the corresponding CCSD/BS2 energies that enter the more sophisticated extrapolation formula given by Eq. (4.28). In analogy to the earlier investigation of bond dissociation curves [291], in applying Eq. (4.29) to extrapolate the CR-CC(2,3)/BS2 reaction pathways, the implicit

assumption is made that the CR-CC(2,3)/BS1 approach is capable of describing the relevant nondynamical correlation effects, which do not change much with the basis set, and the leading dynamical correlations, which can be captured by CR-CC(2,3)/BS1 computations, and that the difference between the UMP2/BS2 and UMP2/BS1 or UMP3/BS2 and UMP3/BS1 energies provides reasonably accurate information about changes in the dynamical correlation effects when going from BS1 to BS2 (the contributions due to nondynamical correlation effects, which UMP2 and UMP3 may describe poorly, but which are almost independent of the basis set, should approximately cancel out when the UMP2/BS2 – UMP2/BS1 or UMP3/BS2 – UMP3/BS1 difference is computed). While the main method used for extrapolating the desired high-level CR-CC(2,3)/BS2 energetics is based on Eq. (4.28), which uses the CCSD/BS1, CCSD/BS2, and CR-CC(2,3)/BS1 data, it is also worth examining whether the considerably less expensive extrapolation approaches based on Eq. (4.29), which are a lot easier to exploit in practice, particularly when larger multiple transition-metal atom systems are to be considered, is similarly effective in the context of examining the mechanism of methanol oxidation to formic acid catalyzed by the  $\text{Au}_8^-$  nanoparticle.

#### 4.3.4 Results and discussion

The results of our calculations for pathway I of the  $\text{CH}_3\text{O}^- + \text{O}_2 + \text{Au}_8^-$  reactive system are presented in Tables 4.5–4.8 and Fig. 4.5. The most essential tables are Tables 4.5 and 4.6, which compare the energy differences for the various species along pathway I resulting from the unrestricted M06, M06-L, BP86, TPSS, B97-D, B3LYP, and  $\omega$ -B97X-D DFT calculations and their restricted CCSD, CR-CC(2,3),A  $\equiv$  CCSD(2)<sub>T</sub>, and CR-CC(2,3),D counterparts



obtained with the ROHF canonicalization schemes of Guest and Saunders. Table 4.5 focuses on the results obtained with the smaller BS1 basis set, whereas Table 4.6 compares the various CC and DFT energies corresponding to the larger BS2 basis. As implied by the remarks above, the CCSD and DFT energies corresponding to the BS2 basis set represent the results of the true calculations with that basis, whereas the CR-CC(2,3),A/BS2 and CR-CC(2,3),D/BS2 energies shown in Table 4.6 are the results of the extrapolations based on Eq. (4.28), in which the larger-basis-set CCSD/BS2 energies have been corrected using the triples contributions extracted from the CR-CC(2,3) calculations using the smaller BS1 basis. Figure 4.5, which accompanies Tables 4.5 and 4.6, shows the reference CR-CC(2,3),D/BS1 and CR-CC(2,3),D/BS2 energetics obtained using the Guest-Saunders ROHF approach, the corresponding M06, B3LYP, and  $\omega$ -B97X-D data, and the molecular structures defining the stationary points along pathway I resulting from the M06 geometry optimizations carried out in Ref. [300]. In addition to the individual energy differences characterizing the various species along pathway I, we show in Tables 4.5 and 4.6 the mean unsigned error (MUE) and nonparallelity error (NPE) values relative to CR-CC(2,3),D calculated using the energies of structures A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, and D1 relative to the reactants. The MUE and NPE values relative to CR-CC(2,3),D determined on the basis of the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences, which represent the three activation energies that are key to understanding the energetics of pathway I and which are less sensitive to the basis set than the differences relative to the isolated reactants, are given in parenthesis in Tables 4.5 and 4.6. The energies of the various species relative to the  $\text{CH}_3\text{O}^-$ ,  $\text{O}_2$ , and  $\text{Au}_8^-$  reactants may also be affected by basis set superposition error (BSSE),

which one cannot determine in a rigorous manner due to the significant rearrangements in the molecular geometries when going from the noninteracting reactants to the stationary points along pathway I. Although the BSSE effects are not expected to be significant when the BS2 basis set is employed, it is useful to base at least part of our error analysis on the (TS\_A1 - A1), (TS\_B1 - B1), and (TS\_C1 - C1) energy differences, which are not affected by BSSE, which vary relatively little when one changes the basis set from BS1 to BS2, and which are physically the most important for understanding the catalytic reaction mechanism represented by pathway I.

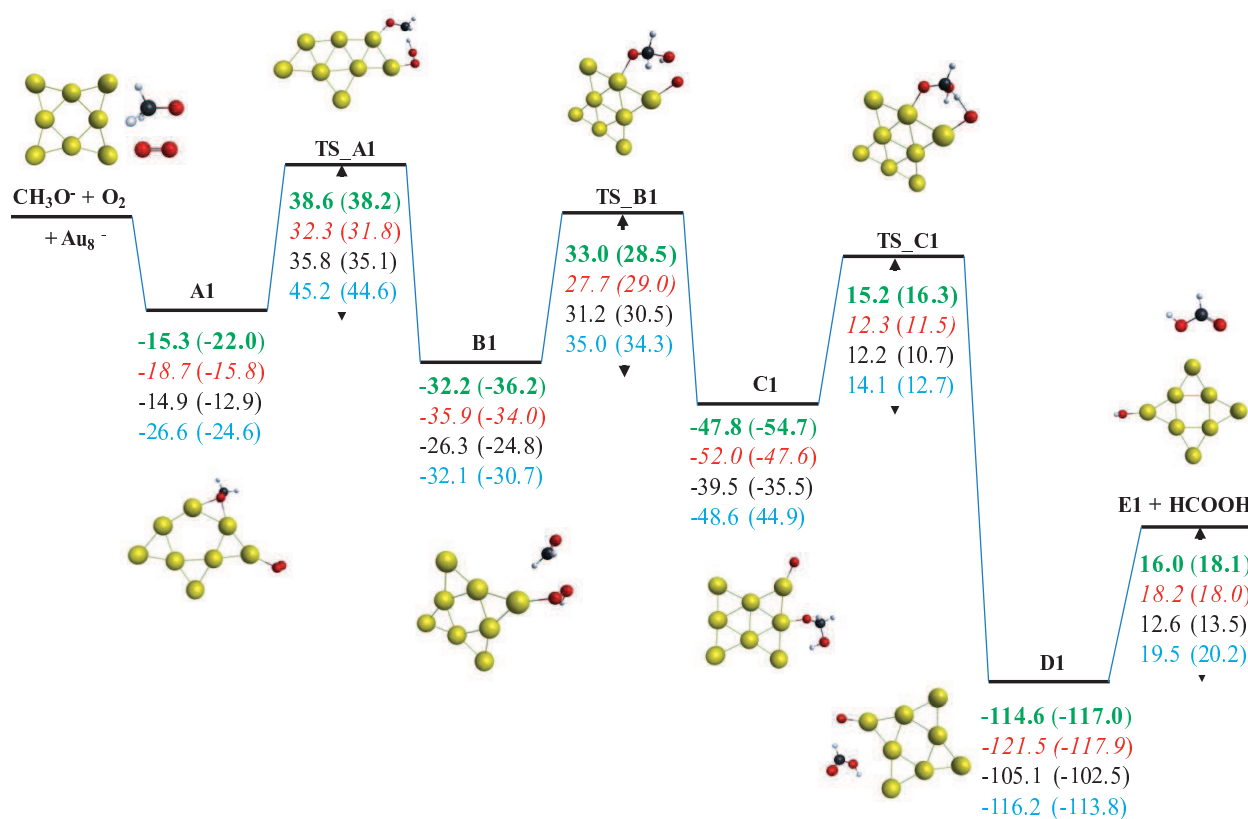


Figure 4.5: Energy diagram for reaction pathway I characterizing the oxidation of methanol to formic acid on the  $\text{Au}_8^-$  particle resulting from CR-CC(2,3),D (green bold font), M06 (red italic font), B3LYP (black roman font), and  $\omega$ -B97X-D (blue roman font) calculations using the BS2 and, in parentheses, BS1 basis sets. The CR-CC(2,3),D/BS1 energies and all of the DFT energies represent truly computed data. The CR-CC(2,3),D/BS2 energies were obtained by extrapolation based on Eq. (4.28), in which the difference between the CCSD energies obtained with the BS2 and BS1 basis sets is added to the CR-CC(2,3),D/BS1 energy. The energies shown for the A1, B1, C1, and D1 intermediates are relative to the  $\text{CH}_3\text{O}^- + \text{O}_2 + \text{Au}_8^-$  reactants. The energies shown for the products are relative to the D1 intermediate. The energies shown for the transition states TS\_A1, TS\_B1, and TS\_C1 are the corresponding (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barrier heights. All energies are in kcal/mol. Also shown are the molecular structures defining the stationary points along pathway I resulting from the M06 geometry optimizations carried out in Ref. [300].

Table 4.5: Energy differences (in kcal/mol) for the various species along pathway I computed using selected CC and DFT approaches and the BS1 basis set.<sup>a</sup>

|                      | CCSD     | CR(2,3),A <sup>b</sup> | CR(2,3),D <sup>b</sup> | M06 <sup>c</sup> | M06-L     | B3LYP     | BP86      | B97-D     | TPSS      | $\omega$ -B97X-D |
|----------------------|----------|------------------------|------------------------|------------------|-----------|-----------|-----------|-----------|-----------|------------------|
| A1 – reactants       | -23.3    | -24.7                  | -22.0                  | -15.8            | -18.9     | -12.9     | -19.6     | -19.2     | -20.7     | -24.6            |
| TS_A1 – reactants    | 21.3     | 14.1                   | 16.1                   | 16.0             | 7.7       | 22.2      | 2.8       | 12.1      | 4.7       | 20.0             |
| TS_A1 – A1           | 44.6     | 38.8                   | 38.2                   | 31.8             | 26.6      | 35.1      | 22.4      | 31.2      | 25.4      | 44.6             |
| B1 – reactants       | -36.1    | -38.7                  | -36.2                  | -34.0            | -29.9     | -24.8     | -29.7     | -29.3     | -28.0     | -30.7            |
| TS_B1 – reactants    | -2.3     | -9.9                   | -7.7                   | -5.0             | -9.6      | 5.6       | -9.9      | -3.6      | -9.1      | 3.6              |
| TS_B1 – B1           | 33.9     | 28.8                   | 28.5                   | 29.0             | 20.3      | 30.5      | 19.8      | 25.7      | 18.9      | 34.3             |
| C1 – reactants       | -54.3    | -57.1                  | -54.7                  | -47.6            | -46.5     | -35.5     | -42.3     | -37.5     | -44.5     | -44.9            |
| TS_C1 – reactants    | -34.9    | -40.4                  | -38.4                  | -36.1            | -38.0     | -24.9     | -38.1     | -33.5     | -36.1     | -32.2            |
| TS_C1 – C1           | 19.4     | 16.8                   | 16.3                   | 11.5             | 8.4       | 10.7      | 4.2       | 4.0       | 8.5       | 12.7             |
| D1 – reactants       | -117.1   | -119.8                 | -117.0                 | -117.9           | -109.8    | -102.5    | -105.5    | -108.3    | -100.9    | -113.8           |
| products – reactants | -98.1    | -99.8                  | -98.9                  | -99.9            | -97.8     | -89.0     | -92.1     | -85.9     | -86.3     | -93.6            |
| products – D1        | 19.0     | 20.0                   | 18.1                   | 18.0             | 18.0      | 13.5      | 13.4      | 22.4      | 14.6      | 20.2             |
| MUE <sup>d</sup>     | 2.1(5.0) | 2.2(0.5)               | 0.0(0.0)               | 2.8(3.9)         | 5.3(9.2)  | 12.1(3.6) | 6.9(12.2) | 7.7(7.4)  | 7.9(10.1) | 6.0(5.3)         |
| NPE <sup>e</sup>     | 6.7(3.3) | 1.9(0.3)               | 0.0(0.0)               | 8.1(6.9)         | 16.6(3.7) | 13.1(7.6) | 25.7(7.1) | 21.2(9.5) | 27.5(5.0) | 13.9(10.0)       |

<sup>a</sup>For the definitions of the various species along pathway I, see Figs. 4.4 and 4.5. The CC results correspond to the Guest-Saunders canonicalization scheme for ROHF orbitals [314].

<sup>b</sup>CR(2,3),A = CR-CC(2,3),A is equivalent to the CCSD(2)<sub>T</sub> approach of Ref. [77]. CR(2,3),D = CR-CC(2,3),D.

<sup>c</sup>Data taken from Ref. [300].

<sup>d</sup>Mean unsigned error relative to CR-CC(2,3),D calculated using the energies of A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, D1, and products relative to the reactants and, in parentheses, using the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences.

<sup>e</sup>Nonparallelity error relative to CR-CC(2,3),D calculated using the energies of A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, D1, and products relative to the reactants and, in parentheses, using the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences.

Table 4.6: Energy differences (in kcal/mol) for the various species along pathway I computed using selected CC and DFT approaches and the BS2 basis set.<sup>a</sup>

|                      | CCSD     | CR(2,3),A <sup>b</sup> | CR(2,3),D <sup>b</sup> | M06 <sup>c</sup> | M06-L     | B3LYP     | BP86      | B97-D     | TPSS      | $\omega$ -B97X-D |
|----------------------|----------|------------------------|------------------------|------------------|-----------|-----------|-----------|-----------|-----------|------------------|
| A1 – reactants       | -16.6    | -18.0                  | -15.3                  | -18.7            | -21.9     | -14.9     | -21.3     | -20.4     | -22.0     | -26.6            |
| TS_A1 – reactants    | 28.5     | 21.3                   | 23.3                   | 13.6             | 5.4       | 20.9      | 1.9       | 11.3      | 3.5       | 18.6             |
| TS_A1 – A1           | 45.0     | 39.3                   | 38.6                   | 32.3             | 27.3      | 35.8      | 23.2      | 31.7      | 25.5      | 45.2             |
| B1 – reactants       | -32.2    | -34.7                  | -32.2                  | -35.9            | -31.8     | -26.3     | -31.0     | -30.1     | -29.4     | -32.1            |
| TS_B1 – reactants    | 6.2      | -1.4                   | 0.8                    | -8.2             | -11.3     | 4.9       | -11.0     | -4.0      | -10.3     | 2.9              |
| TS_B1 – B1           | 38.4     | 33.3                   | 33.0                   | 27.2             | 20.5      | 31.2      | 19.9      | 26.1      | 19.1      | 35.0             |
| C1 – reactants       | -47.4    | -50.3                  | -47.8                  | -52.0            | -51.0     | -39.5     | -47.1     | -41.6     | -49.0     | -48.6            |
| TS_C1 – reactants    | -29.1    | -34.6                  | -32.6                  | -39.7            | -40.8     | -27.3     | -41.0     | -35.7     | -39.0     | -34.5            |
| TS_C1 – C1           | 18.3     | 15.7                   | 15.2                   | 12.3             | 10.2      | 12.2      | 6.1       | 5.8       | 10.1      | 14.1             |
| D1 – reactants       | -114.8   | -117.4                 | -114.6                 | -121.5           | -112.2    | -105.1    | -108.2    | -110.4    | -103.6    | -116.2           |
| products – reactants | -97.9    | -99.6                  | -98.6                  | -103.4           | -94.2     | -92.5     | -95.8     | -89.2     | -90.0     | -96.7            |
| products – D1        | 16.9     | 17.9                   | 16.0                   | 18.2             | 18.0      | 12.6      | 12.4      | 21.2      | 13.6      | 19.5             |
| MUE <sup>d</sup>     | 2.1(5.0) | 2.2(0.5)               | 0.0(0.0)               | 6.1(4.9)         | 6.9(9.6)  | 5.2(2.6)  | 7.4(12.5) | 5.9(7.7)  | 8.5(10.7) | 3.1(3.2)         |
| NPE <sup>e</sup>     | 6.7(3.3) | 1.9(0.3)               | 0.0(0.0)               | 6.3(3.4)         | 22.3(7.4) | 11.9(1.2) | 27.8(6.3) | 21.4(2.4) | 30.8(8.8) | 13.5(7.7)        |

<sup>a</sup>For the definitions of the various species along pathway I, see Figs. 4.4 and 4.5. The CC results correspond to the Guest-Saunders canonicalization scheme for ROHF orbitals [314].

<sup>b</sup>CR(2,3),A = CR-CC(2,3),A is equivalent to the CCSD(2)<sub>T</sub> approach of Ref. [77]. CR(2,3),D = CR-CC(2,3),D. Extrapolated using Eq. (4.28).

<sup>c</sup>Data taken from Ref. [300].

<sup>d</sup>Mean unsigned error relative to CR-CC(2,3),D calculated using the energies of A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, D1, and products relative to the reactants and, in parentheses, using the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences.

<sup>e</sup>Nonparallelity error relative to CR-CC(2,3),D calculated using the energies of A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, D1, and products relative to the reactants and, in parentheses, using the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences.

Table 4.7 compares the CCSD/BS1, CR-CC(2,3),A/BS1, and CR-CC(2,3),D/BS1 results obtained with the two types of canonical ROHF orbitals considered in this study. Examining Table 4.7 it is immediately obvious that for each CC method analyzed in this work, the impact of the different ROHF canonicalization schemes on the resulting energetics is negligible, as the differences between the CCSD, CR-CC(2,3),A, and CR-CC(2,3),D energies obtained with the Roothaan and Guest-Saunders approaches do not exceed 0.1, 0.2, and 0.3 kcal/mol, respectively. Thus, in the rest of the discussion in this section, the CC results based on the frequently exploited Guest-Saunders canonicalization approach, summarized in Tables 4.5 and 4.6, will be used.

Table 4.8 compares the energy differences for the various species along pathway I resulting from the two different ways of extrapolating the larger-basis-set CC/BS2 energetics discussed in Section 4.3.3 and represented by Eqs. (4.28) and (4.29). Although much of the discussion below focuses on the more accurate procedure for extrapolating the CR-CC(2,3)/BS2 energetics based on Eq. (4.28), in which the bulk of the electron correlation energy captured by CCSD is determined using the target BS2 basis, reducing the use of BS1 to the triples corrections only, it is interesting to observe that one can obtain similar results by replacing CCSD in the basis set extrapolation scheme by the much less expensive UMP2 or UMP3 approach, eliminating the need to carry out the larger-basis-set CC calculations altogether. As shown in Table 4.8, the energy differences for the various species along pathway I resulting from the true CCSD/BS2 calculations and their extrapolated counterparts using Eq. (4.29) agree to within  $\sim 3$  kcal/mol in the UMP2 case and  $\sim 2$  kcal/mol when the UMP3 energies are used, though there are several cases where the true CCSD/BS2 results and the

Table 4.7: Energy differences (in kcal/mol) for the various species along pathway I computed at the CCSD and CR-CC(2,3) levels using the BS1 basis set and two different schemes for obtaining the canonical ROHF orbitals.

|                      | CCSD           |                 | CR-CC(2,3),A <sup>a</sup> |                 | CR-CC(2,3),D   |                 |
|----------------------|----------------|-----------------|---------------------------|-----------------|----------------|-----------------|
|                      | R <sup>b</sup> | GS <sup>c</sup> | R <sup>b</sup>            | GS <sup>c</sup> | R <sup>b</sup> | GS <sup>c</sup> |
| A1 – reactants       | -23.3          | -23.3           | -24.4                     | -24.7           | -21.9          | -22.0           |
| TS_A1 – reactants    | 21.3           | 21.3            | 14.3                      | 14.1            | 16.4           | 16.1            |
| TS_A1 – A1           | 44.6           | 44.6            | 38.7                      | 38.8            | 38.3           | 38.2            |
| B1 – reactants       | -36.1          | -36.1           | -38.6                     | -38.7           | -36.1          | -36.2           |
| TS_B1 – reactants    | -2.2           | -2.3            | -9.8                      | -9.9            | -7.7           | -7.7            |
| TS_B1 – B1           | 33.9           | 33.9            | 28.8                      | 28.8            | 28.4           | 28.5            |
| C1 – reactants       | -54.3          | -54.3           | -57.0                     | -57.1           | -54.5          | -54.7           |
| TS_C1 – reactants    | -34.9          | -34.9           | -40.2                     | -40.4           | -38.3          | -38.4           |
| TS_C1 – C1           | 19.4           | 19.4            | 16.7                      | 16.8            | 16.2           | 16.3            |
| D1 – reactants       | -117.1         | -117.1          | -119.7                    | -119.8          | -117.1         | -117.0          |
| products – reactants | -98.1          | -98.1           | -99.7                     | -99.8           | -98.7          | -98.9           |
| products – D1        | 19.0           | 19.0            | 20.0                      | 20.0            | 18.3           | 18.1            |

<sup>a</sup>Equivalent to the CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>b</sup>Results obtained using the Roothaan canonicalization scheme for ROHF orbitals [313].

<sup>c</sup>Results obtained using the Guest-Saunders canonicalization scheme for ROHF orbitals [314].

UMP3-extrapolated energetics agree to within less than 0.5 kcal/mol. In most cases, including, for example, the rate-determining (TS\_A1 – A1) barrier, which measures the amount of energy needed to go from the O<sub>2</sub> and methoxy molecules coadsorbed on the Au<sub>8</sub><sup>–</sup> particle to the initial transition state TS\_A1, and the overall reaction energy, the agreement is even better, with the differences between the true and UMP $n$ -based extrapolated CCSD/BS2 data oscillating around 1–2 kcal/mol or less. Since we currently have no access to the true CR-CC(2,3)/BS2 data, we cannot make analogous comparisons, but, in view of the fact that the connected triply excited clusters account for a small fraction of the correlation effects relative to the CCSD contributions, it is rather safe to assume that similar levels of accuracy apply to the CR-CC(2,3)/BS2 energies extrapolated using Eq. (4.29), which are presented in Table 4.8 as well. The relatively small differences between the true CCSD/BS2 energies and their UMP $n$ -based extrapolated counterparts obtained using Eq. (4.29) automatically guarantee that the same small differences must be observed when comparing the CR-CC(2,3),A/BS2 and CR-CC(2,3),D/BS2 energetics extrapolated using Eq. (4.29) with their counterparts determined using the extrapolation scheme defined by Eq. (4.28).



Table 4.8: Comparisons of energy differences (in kcal/mol) for the various species along pathway I obtained using two different ways of extrapolating the CC/BS2 energetics,<sup>a</sup> represented by Eqs. (4.28) and (4.29).

|                      | CCSD              |                   |                   | CR-CC(2,3),A <sup>b</sup> |                   |                   | CR-CC(2,3),D      |                   |                   |
|----------------------|-------------------|-------------------|-------------------|---------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                      | CCSD <sup>c</sup> | UMP2 <sup>d</sup> | UMP3 <sup>d</sup> | CCSD <sup>e</sup>         | UMP2 <sup>d</sup> | UMP3 <sup>d</sup> | CCSD <sup>e</sup> | UMP2 <sup>d</sup> | UMP3 <sup>d</sup> |
| A1 – reactants       | -16.6             | -13.6             | -16.9             | -18.0                     | -15.0             | -18.3             | -15.3             | -12.3             | -15.7             |
| TS_A1 – reactants    | 28.5              | 31.7              | 30.5              | 21.3                      | 24.5              | 23.3              | 23.3              | 26.5              | 25.3              |
| TS_A1 – A1           | 45.0              | 45.3              | 47.4              | 39.3                      | 39.5              | 41.6              | 38.6              | 38.8              | 41.0              |
| B1 – reactants       | -32.2             | -32.6             | -30.8             | -34.7                     | -35.1             | -33.3             | -32.2             | -32.6             | -30.8             |
| TS_B1 – reactants    | 6.2               | 4.1               | 7.6               | -1.4                      | -3.5              | 0.0               | 0.8               | -1.3              | 2.2               |
| TS_B1 – B1           | 38.4              | 36.7              | 38.4              | 33.3                      | 31.7              | 33.3              | 33.0              | 31.3              | 33.0              |
| C1 – reactants       | -47.4             | -51.6             | -45.4             | -50.3                     | -54.5             | -48.2             | -47.8             | -52.1             | -45.8             |
| TS_C1 – reactants    | -29.1             | -31.0             | -27.1             | -34.6                     | -36.5             | -32.5             | -32.6             | -34.5             | -30.6             |
| TS_C1 – C1           | 18.3              | 20.6              | 18.3              | 15.7                      | 18.0              | 15.7              | 15.2              | 17.5              | 15.2              |
| D1 – reactants       | -114.8            | -117.1            | -112.8            | -117.4                    | -119.7            | -115.5            | -114.6            | -116.9            | -112.7            |
| products – reactants | -97.9             | -98.9             | -97.0             | -99.6                     | -100.6            | -98.7             | -98.6             | -99.7             | -97.8             |
| products – D1        | 16.9              | 18.1              | 15.8              | 17.9                      | 19.1              | 16.7              | 16.0              | 17.3              | 14.9 <sup>c</sup> |

<sup>a</sup>All of the CC energies used in the basis set extrapolations correspond to the Guest-Saunders canonicalization scheme for ROHF orbitals [314].

<sup>b</sup>Equivalent to the CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>c</sup>Results of the true CCSD/BS2 calculations.

<sup>d</sup>Extrapolated by adding the difference between the UMP<sub>n</sub> ( $n = 2, 3$ ) energies obtained with the BS2 and BS1 basis sets to the appropriate CC/BS1 energy, as in Eq. (4.29).

<sup>e</sup>Extrapolated by adding the difference between the CCSD energies obtained with the BS2 and BS1 basis sets to the appropriate CR-CC(2,3)/BS1 energy, as in Eq. (4.28).

Clearly, if the larger-basis-set CCSD computations are affordable, as is the case in this study, where the aerobic oxidation of methanol catalyzed by the  $\text{Au}_8^-$  nanocluster is examined, the extrapolation of the CR-CC(2,3)/BS2 results that combines the CCSD/BS1, CCSD/BS2, and CR-CC(2,3)/BS1 data (Eq. (4.28)) is the recommended approach. However, one may encounter difficulties with performing larger-basis-set CCSD calculations for the analogous reactions catalyzed by larger  $\text{Au}_n^-$  particles, such as  $\text{Au}_{20}^-$ , which was investigated using DFT in Ref. [299]. Using the  $\text{MP}n$ -based ( $n = 2, 3$ ) extrapolation, as in Eq. (4.29), to obtain the larger-basis-set CR-CC(2,3) information, where the effects of the one-electron basis on the CR-CC(2,3) energies are extracted from the relatively inexpensive UMP2 or UMP3 calculations, may offer a good alternative to the procedure based on Eq. (4.28). Since we have access to the more accurate CR-CC(2,3)/BS2-level data extrapolated using the more accurate Eq. (4.28), we rely on this equation in the remaining discussion, though, one should keep in mind that the CR-CC(2,3)/BS2-level information obtained using the less demanding Eq. (4.29) is not much different.

#### 4.3.4.1 Coupled-cluster calculations

According to the CCSD, CR-CC(2,3),A, and CR-CC(2,3),D methods and in agreement with the experimental [202] and previously obtained theoretical [299, 300] data, the reaction examined here is exothermic and the conversion from the methoxy species to formaldehyde via transition state TS\_A1 is the rate determining step, independent of the basis set employed in the calculations. As shown in Table 4.6, our best CR-CC(2,3),D/BS2 estimates, obtained using Eq. (4.28), place the initial transition state TS\_A1, corresponding to the hydrogen transfer from the methoxy species to the molecular oxygen to form intermediate B1, at

23.3 kcal/mol above the isolated reactants (38.6 kcal/mol above structure A1 representing  $\text{O}_2$  and  $\text{CH}_3\text{O}^-$  coadsorbed on  $\text{Au}_8^-$ ) and considerably above the remaining two transition states characterizing pathway I, namely, 22.5 kcal/mol above TS\_B1 and 55.9 kcal/mol above TS\_C1. As shown in Table 4.5, the use of the smaller BS1 basis lowers the initial barrier relative to the isolated reactants to 16.1 kcal/mol, most likely because of the larger BSSE, but the  $(\text{TS\_A1} - \text{TS\_B1})$ , and  $(\text{TS\_A1} - \text{TS\_C1})$  differences of 23.8 and 54.5 kcal/mol, respectively, resulting from the CR-CC(2,3),D/BS1 calculations, which are not affected by BSSE, are almost the same as in the BS2-based computations. We can, thus, conclude, with great deal of confidence, that after passing the initial barrier defined by TS\_A1, the oxidation of methanol to formic acid catalyzed by the  $\text{Au}_8^-$  particle is energetically an overall “down-hill” process that proceeds via two additional barriers located considerably below TS\_A1, resulting in the products being located almost 100 kcal/mol below the reactants. All three CC theories used for this study provide the same picture in this regard.

While the energies of the various species defining pathway I relative to the reactants obtained with the CC approaches are somewhat sensitive to the basis set, with differences between the estimated BS2 and calculated BS1 data ranging from less than 1 kcal/mol for the products to  $\sim 8$  kcal/mol for the TS\_B1 transition state (mostly due to the BSSE effects, which affect the energies of the isolated  $\text{CH}_3\text{O}^-$ ,  $\text{O}_2$ , and  $\text{Au}_8^-$  reactants, particularly in the calculations employing the smaller BS1 basis), the key activation barriers, defined as the  $(\text{TS\_A1} - \text{A1})$ ,  $(\text{TS\_B1} - \text{B1})$ , and  $(\text{TS\_C1} - \text{C1})$  energy differences, especially  $(\text{TS\_A1} - \text{A1})$ , and the overall reaction energy depend relatively little on the basis set. For example, the CR-CC(2,3),D/BS2 estimates of the  $(\text{TS\_A1} - \text{A1})$ ,  $(\text{TS\_B1} - \text{B1})$ , and  $(\text{TS\_C1} -$

C1) barrier heights and the overall reaction energy based on Eq. (4.28) are 38.6, 33.0, 15.2, and -98.6 kcal/mol, respectively. These should be compared to the results of the CR-CC(2,3),D/BS1 calculations, which give 38.2, 28.5, 16.3, and -98.9 kcal/mol, respectively, for the same quantities, in overall good agreement with the CR-CC(2,3),D/BS2 estimates. Even the more primitive estimates of the CR-CC(2,3),D/BS2 energetics relying on the basis set dependence extracted from the UMP2/BS1 and UMP2/BS2 or the UMP3/BS1 and UMP3/BS2 calculations, Eq. (4.29), give similar values of the (TS\_A1 - A1), (TS\_B1 - B1), and (TS\_C1 - C1) energy differences and the reaction energy (38.8, 31.3, 17.5, and -99.7 kcal/mol, respectively, in the UMP2-based case and 41.0, 33.0, 15.2, and -97.8 kcal/mol, respectively, when using information from the UMP3 calculations; see Table 4.8).

All of this implies that the CR-CC(2,3),D values of the (TS\_A1 - A1), (TS\_B1 - B1), and (TS\_C1 - C1) barriers, which are not affected by BSSE and which do not change much when we switch from the smaller BS1 basis set to the larger BS2 basis, especially in the (TS\_A1 - A1) and (TS\_C1 - C1) cases, and the corresponding reaction energy, which displays a weak basis set dependence as well, are particularly useful for judging the various DFT functionals considered in this study. Some other energy differences along pathway I listed in Tables 4.5 and 4.6 may be somewhat less robust because of their stronger basis set dependence, but the sophisticated *ab initio* level of theory represented in this work by the CR-CC(2,3),D/BS2 approach, where we use a rather large basis set in the underlying CCSD calculations and the advanced CR-CC(2,3) treatment of the high-order correlation effects beyond CCSD, is certainly high enough in the present case to benchmark DFT, where the variation in the results obtained with various functionals is much greater than that obtained

in the CR-CC(2,3) considerations. As pointed out above and as shown in numerous earlier applications and benchmark studies, such as those reported in Refs. [51,93–95,98,99,105,290–292,323,324], the CR-CC(2,3) methodology, particularly its variant D, is capable of providing results in the chemical accuracy ( $\sim 1\text{--}2$  kcal/mol) range for activation, reaction, and binding energies when basis sets of triple- $\zeta$  quality with polarization and diffuse functions, such as BS2, are employed. Obviously the CR-CC(2,3) results employing the smaller BS1 basis are not as accurate and may carry larger errors, but they can still be useful in assessing trends in DFT. Indeed, in agreement with the earlier test applications involving CR-CC(2,3), such as those presented in Refs. [98, 99, 323, 324], the average unsigned difference between the triple- $\zeta$ -level CR-CC(2,3)/BS2 and double- $\zeta$ -level CR-CC(2,3)/BS1 data (which in our case is the same as the average unsigned difference between the CCSD/BS2 and CCSD/BS1 results because of the use of Eq. (4.28)) obtained using the energies of A1, TS\_A1, B1, TS\_B1, C1, TS\_C1, D1, and products relative to the reactants is about 5 kcal/mol. The analogous difference between the CR-CC(2,3)/BS2 and CR-CC(2,3)/BS1 results based on the most important (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barriers is only 2 kcal/mol. This means that even the CR-CC(2,3)/BS1 calculations, which are certainly not converged with respect to the one-electron basis set, but still offer a well-balanced treatment of dynamical and nondynamical correlation effects that the lower-order methods based on DFT do not always describe, may be quite valuable. Tables 4.5 and 4.6 indeed show that this is the case, with most of the DFT functionals producing errors in the relative energies and variations in the results for almost every structure along pathway I that exceed by a wide margin the 2–5 kcal/mol average basis set errors characterizing the CR-CC(2,3)/BS1

data. However, in examining the DFT data, we will rely on our best CR-CC(2,3),D/BS2 energetics, which uses a larger basis set and carries smaller errors, to ensure our conclusions are more definitive.

As can be seen in Table 4.5, the MUE values characterizing the CCSD and CR-CC(2,3),A calculations relative to CR-CC(2,3),D, which are on the order of 2 kcal/mol, are similar, supporting our earlier assertion that the CR-CC(2,3),D results can be regarded as reasonably well converged with the many-electron correlation effects in a given basis set. However, as is generally the case, the CCSD approach alone is not sufficiently accurate to provide a reliable description of the activation barriers. Indeed, the differences between the CR-CC(2,3),D/BS1 and CCSD/BS1 activation energies characterizing pathway I range from about 3 to 6 kcal/mol, with the largest difference being obtained for TS\_A1 and TS\_B1 ( $\sim 5 - 6$  kcal/mol). Furthermore, as observed in other applications, the CCSD approach, which neglects connected triple excitations, overestimates the activation energies. It is sufficient to use CCSD to obtain an accurate description of the A1, B1, and C1 intermediates and the overall reaction energy (the differences between CR-CC(2,3),D and CCSD are in this case on the order of 1 kcal/mol or smaller), but one needs to incorporate connected triples in the CC calculations to obtain a consistently accurate description of the entire reaction pathway. This is seen by comparing the CR-CC(2,3),A and CR-CC(2,3),D data calculated with the BS1 basis set, summarized in Table 4.5. Although the MUE values characterizing the CCSD and CR-CC(2,3),A computations relative to CR-CC(2,3),D are similar, there is a great deal of consistency between the individual CR-CC(2,3),A and CR-CC(2,3),D energies, differing by about 1–2 kcal/mol at all stationary points. When we replace CR-CC(2,3),A by CCSD,

comparing with CR-CC(2,3),D, this is no longer the case and, as already mentioned, the agreement is no longer as good.

#### 4.3.4.2 Benchmarking DFT against the reference CR-CC(2,3),D data

We now turn to the comparison of our best CR-CC(2,3),D results with those obtained from the various DFT functionals, focusing on the results obtained with the larger BS2 basis set, as summarized in Table 4.6 and Fig. 4.5. In our published work [181], we concluded that among the several representative DFT methods considered in Ref. [181], the M06 approach used in the earlier calculations for the same reaction pathway [300] and the highly popular hybrid B3LYP functional are generally most accurate. Indeed, the overall MUE and NPE values relative to CR-CC(2,3),D/BS2 using all stationary points along pathway I are 6.1 and 6.3 kcal/mol, respectively, in the M06 case and 5.2 and 11.9 kcal/mol, respectively, for B3LYP which is a better performance compared to all other DFT calculations included in Ref. [181] (i.e., all listed in Tables 4.5 and 4.6 other than  $\omega$ -B97X-D). With the additional and previously unpublished  $\omega$ -B97X-D results included in this dissertation, the situation seems to be changing, since the MUE and NPE relative to CR-CC(2,3),D/BS2 characterizing  $\omega$ -B97X-D of 3.1 and 13.5 kcal/mol, respectively, look competitive, when compared to M06 and B3LYP, but this is only the initial impression. While the MUE for the  $\omega$ -B97X-D results appears to be better than those for the M06 and B3LYP approaches, with an NPE larger than NPEs characterizing M06 and B3LYP and close to 14 kcal/mol, it is hard to make the claim that  $\omega$ -B97X-D performs better. Obviously, none of the DFT methods look good compared with, for example, the CCSD MUE value of 2.1 kcal/mol, but the M06 and, especially, B3LYP functionals perform substantially better than those characterizing the remaining

DFT calculations presented in Table 4.6. Indeed, both hybrid functionals, M06 and B3LYP, particularly the latter, perform well in describing the key (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barrier heights, which are 35.8, 31.2, and 12.2 kcal/mol, respectively, when the B3LYP/BS2 approach is used, and 32.3, 27.7, and 12.3 kcal/mol, respectively, when the M06/BS2 method is employed, in reasonably a nice agreement with the corresponding CR-CC(2,3),D/BS2 benchmark data, which are 38.6, 33.0, and 15.2 kcal/mol, respectively, for the same three barrier heights. This is reflected in the considerable improvements in the MUE and NPE values relative to CR-CC(2,3),D/BS2 resulting from the corresponding B3LYP and M06 calculations, which are reduced to 2.6 and 1.2 kcal/mol, respectively, in the former case, and 4.9 and 3.4 kcal/mol, respectively, in the latter case when we limit ourselves to the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) activation energies. The B3LYP description of these three barriers, which is better than that provided by CCSD, is especially encouraging, placing the initial (TS\_A1 – A1) barrier at 35.8 kcal/mol, that is only 2.8 kcal/mol apart from our best CR-CC(2,3),D/BS2 estimate; the CCSD approach overestimates the same barrier by about 6 kcal/mol and M06 underestimates it by more or less the same amount. The previously unpublished  $\omega$ -B97X-D/BS2 results for the three barrier heights, (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1), are 45.2, 35.0, and 14.1 kcal/mol, respectively, i.e., they are not nearly as good as those obtained with B3LYP. As such, we do not include this functional in the rest of this discussion and focus on the functionals included in our published study [181].

There is also good agreement between the B3LYP/BS2 and CR-CC(2,3),D/BS2 energies of transition states TS\_A1, TS\_B1, and TS\_C1 relative to the isolated reactants, which are



20.9, 4.9, and  $-27.3$  kcal/mol, respectively, in the former case and 23.3, 0.8, and  $-32.6$  kcal/mol, respectively, in the case of the latter approach, indicating that the B3LYP functional is capable of providing a reliable description of the bond rearrangements involved in catalytic reactions on metallic nanoparticles of the type represented by the oxidation process catalyzed by gold clusters examined in this study. It looks as though the energy of the TS\_B1 transition state relative to the reactants resulting from the B3LYP/BS2 calculations is in larger error, but we must keep in mind that B3LYP is the only functional in Table 4.6, other than  $\omega$ -B97X-D not included in the our published study [181], and which incapable of properly characterizing the key (TS\_A1 – A1) barrier that produces a small positive barrier in this case, with the CR-CC(2,3),D/BS2 approach doing the same. M06 does not seem to work well in this case, but we should not read too much into it, since the M06/BS2 result for the physically more significant (TS\_B1 – B1) energy difference (27.7 kcal/mol) is still in reasonable agreement with the CR-CC(2,3),D/BS2 result (33.0 kcal/mol). Of all functionals tested here, the only one that can improve the overall pathway I description by M06 is B3LYP.

Interestingly, our finding that the B3LYP and M06 functionals provide the best description of the key activation barriers represented by the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) energy differences remains true when we compare the DFT and CR-CC(2,3),D results obtained with the smaller BS1 basis set, shown in Table 4.5. This is a consequence of (1) the well-established fact that DFT converges fast with respect to the basis set and (2) the previously discussed observation that the CC values of these three energy differences are essentially free from BSSE and largely insensitive to the size of the basis employed in

the CC calculations. As pointed out above and shown in Table 4.8, one can successfully correct the results of the small-basis-set CR-CC(2,3) calculations by carrying out additional UMP2 or UMP3 calculations with a larger basis set and by extrapolating the larger-basis-set CR-CC(2,3)-level data using Eq. (4.29) without the need to perform any larger-basis-set CC computations. These remarks may be useful in future work on catalytic systems larger than the one examined here, where larger-basis-set CC calculations may no longer be feasible, but a combination of smaller-basis-set CR-CC(2,3) and larger-basis-set UMP $n$  ( $n = 2, 3$ ) computations, run, for example, on multiple cores, may still be manageable, allowing one to test various DFT approaches in a very meaningful manner.

Considering the generally good agreement between the CR-CC(2,3),D, B3LYP, and M06 energetics characterizing pathway I, we can certainly conclude that the previously reported M06 calculations on the mechanism of methanol oxidation to formic acid on the Au $_8^-$  and Au $_{20}^-$  particles [299, 300] can be regarded as quite reliable, although it would be desirable to repeat the calculations for the larger catalytic system involving Au $_{20}^-$  using the B3LYP approach, which has here been shown to improve the results for the activation barriers characterizing the analogous Au $_8^-$ -containing system. Unfortunately, none of the other DFT functionals examined in the present study offer satisfactory performance. For example, although the pure GGA-type analog of M06, denoted as M06-L, is characterized by an overall MUE value relative to CR-CC(2,3),D of 6.9 kcal/mol, which is only slightly worse than that obtained with M06 when the BS2 basis is employed, it lowers the activation energies relative to M06 by 6 kcal/mol when the (TS\_A1 – A1) energy difference is considered and 7.2 kcal/mol in the (TS\_B1 – B1) case. As a result, the MUE value relative to CR-CC(2,3),D

characterizing the M06-L calculations for the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barriers increase by factors of 2 relative to M06 when the BS2 basis set is employed and more than 2 in the case of the BS1 basis. The overall NPE value relative to CR-CC(2,3),D based on all of the stationary points involved in pathway I characterizing the M06-L calculations with the BS2 basis set, of 22.3 kcal/mol, represents an increase by a factor of 3.5 relative to M06 and 2 relative to B3LYP. Similar remarks, as discussed above, apply to the long-range corrected hybrid  $\omega$ -B97X-D results, whose MUE is smaller than that for both B3LYP and M06 when the BS2 basis set is used, but its NPE values are substantially larger than those characterizing M06 and B3LYP, especially when we focus on the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) activation barriers.

Similar behavior is observed when other pure functionals are considered. The overall MUE values relative to CR-CC(2,3),D/BS2 characterizing the BP86/BS2 and TPSS/BS2 calculations, of 7.4 and 8.5 kcal/mol, respectively, do not appear to be much higher than the M06 result (6.1 kcal/mol), but this is misleading, since the corresponding NPE values based on all stationary points along pathway I are 27.8 and 30.8 kcal/mol, respectively, indicating a very inaccurate representation of pathway I by the BP86 and TPSS functionals. This is again particularly true when we look at the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barriers resulting from the BP86 and TPSS calculations, which are underestimated by  $\sim 13$ – $15$  kcal/mol in the case of the (TS\_A1 – A1) and (TS\_B1 – B1) differences and about 5 kcal/mol (TPSS) or 9 kcal/mol (BP86) when the (TS\_C1 – C1) energy difference is examined. As a result, the MUE values relative to CR-CC(2,3),D characterizing the BP86 and TPSS calculations for the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) activation

barriers increase to 12.5 and 10.7 kcal/mol, respectively, when the BS2 basis is employed (12.2 and 10.1 kcal/mol, respectively, when one uses BS1). Similar trends are observed when we look at the energies of the transition states TS\_A1, TS\_B1, and TS\_C1 relative to the isolated reactants, which are much too low as well. For example, the activation barriers relative to the reactants corresponding to the transition state TS\_A1, which defines the rate-determining step involving the conversion of the methoxy species to formaldehyde, are 1.9 and 3.5 kcal/mol for the BP86 and TPSS calculations, respectively, when the BS2 basis set is employed. These should be compared with the values of 23.3 kcal/mol obtained in the CR-CC(2,3),D/BS2 calculations and 20.9 kcal/mol obtained with B3LYP/BS2.

One can improve the above results obtained with pure GGAs through the use of the popular Grimme’s empirical dispersion corrections, as in the B97-D case, where the (TS\_A1 – reactants) and (TS\_A1 – A1) energy differences increase from 1.9 and 23.2 kcal/mol when the BP86/BS2 method is used to 11.3 and 31.7 kcal/mol, respectively, when the B97-D/BS2 approach is exploited, bringing the results close to the M06 level, but the overall description of pathway I by B97-D is not as good as that provided by the hybrid M06 and B3LYP functionals. For example, while improving the (TS\_A1 – A1) and (TS\_B1 – B1) barriers resulting from the BP86 calculations, the B97-D functional is incapable of changing the poor BP86 results for the (TS\_C1 – C1) energy difference. B97-D also worsens the overall reaction energy obtained with BP86, which is –95.8 kcal/mol according to the BP86/BS2 calculations and –89.2 kcal/mol according to the corresponding B97-D computations, where the CR-CC(2,3),D approach gives –98.6 kcal/mol when the BS2 basis set is employed. The overall NPE value relative to CR-CC(2,3),D/BS2 characterizing the B97-D/BS2 calculations,

of 21.4 kcal/mol, is almost as bad as that obtained with the M06-L scheme, although one observes an improvement compared to the pure BP86 functional, whose corresponding NPE is 27.8 kcal/mol. The same holds true when we compare the MUE values characterizing the B97-D description of the (TS\_A1 – A1), (TS\_B1 – B1), and (TS\_C1 – C1) barriers with the BS1 and BS2 basis sets, which are better than those provided by BP86, but not as good as the results of the M06 and B3LYP calculations.

In summary, we have shown that our highest-level CR-CC(2,3) calculations confirmed the earlier proposals [299,300] that the oxidation of methanol to formic acid on  $\text{Au}_8^-$  proceeds exothermically and that the rate determining step for the reaction is the initial conversion of the methoxy species to formaldehyde. The initial rate-determining transition state, which corresponds to hydrogen transfer from the methoxy species to the molecular oxygen, is placed about 20 kcal/mol above the reactants, less than 40 kcal/mol above the  $\text{O}_2$  and  $\text{CH}_3\text{O}^-$  species coadsorbed on  $\text{Au}_8^-$ , and considerably above the remaining two transition states along the reaction pathway. The previously exploited [299,300] M06 hybrid functional shows reasonable agreement with CR-CC(2,3), but B3LYP improves the description of the activation barriers compared with the M06 approach, bringing the MUE and NPE values relative to our best CR-CC(2,3),D/BS2 estimates characterizing the key activation barriers from 4.9 and 3.4 kcal/mol, respectively, when the M06 method is employed to 2.6 and 1.2 kcal/mol, respectively, when one uses B3LYP. Clearly, examining methanol oxidation to formic acid on  $\text{Au}_{20}^-$  using CR-CC(2,3) by the approach in which the larger-basis-set CR-CC(2,3) data are extrapolated by combining the smaller-basis-set CR-CC(2,3) results with the larger-basis-set CCSD, UMP2, or UMP3 information, as has been done here, or by taking

advantage of the recently developed multilevel local correlation CR-CC(2,3) methodology [101], which is applicable to larger reactive systems containing transition-metal atoms [99], would be helpful too. We hope the information provided herein will be useful for future theoretical studies of reactive processes catalyzed by transition-metal nanoparticles.

## 4.4 Coupled-Cluster and Multireference Configuration Interaction Studies of the Low-Lying Electronic States of 1,2,3,4-Cyclobutanetetraone

### 4.4.1 Background information and scope of the work

1,2,3,4-cyclobutanetetraone ( $C_4O_4$ ) is a small, but surprisingly complex molecule that poses several challenges for theory and experiment. The early theoretical suggestion that this  $D_{4h}$ -symmetric species has a triplet ground state of the  $B_{2u}$  symmetry, with a closed-shell singlet of the  $A_{1g}$  symmetry located only a few kcal/mol higher [325–328], has recently been strengthened through photodetachment experiments [329, 330]. There have, however, been serious problems in obtaining this result computationally, since even the most sophisticated electronic structure treatments encounter significant difficulties with determining whether the ground state of  $C_4O_4$  is a triplet with nine  $\pi$  electrons or a singlet with eight or ten  $\pi$  electrons [325–327, 331] (cf. Fig. 4.6 for the configurations defining the four lowest-energy electronic states of  $C_4O_4$ ).

Molecules with small gaps between the lowest singlet and triplet states have historically

been a challenge for both experimental and theoretical methods. For example, the methylene biradical, which is one of the most celebrated cases in this area, caused much controversy, especially in the 1960s and 1970s, due not only to the smallness of its singlet-triplet gap, but also because of the direct challenge by various theoretical approaches [332–334] to the initial Herzberg’s experimental findings [335] suggesting that the triplet ground state had a linear geometry (cf. Refs. [177, 178, 192, 271, 272, 336–353] for selected examples of *ab initio* calculations of the ground and excited states of methylene, which is certainly not complete nor exhaustive; see, also, Refs. [354, 355] for personal accounts by Schaefer and Harrison, who made pioneering contributions to computational studies of methylene). Thus, even for tiny molecules, such as CH<sub>2</sub>, determining the ordering of close- and low-lying electronic states can be a challenge for single- and multireference quantum chemical approaches. C<sub>4</sub>O<sub>4</sub>, which has many more electrons and eight non-hydrogen atoms and which is related to other similarly challenging larger organic molecules, such as tetrakis-annelated cyclooctatetraene that has been predicted to have a small gap between the lowest singlet and triplet electronic states [356], has a tiny singlet-triplet gap smaller than that in methylene, while having near-degenerate HOMO and LUMO, resulting in two additional singlet states which are also very close in energy. Thus, there are four low-lying nearly degenerate electronic states in 1,2,3,4-cyclobutanetetraone that one has to deal with, which is a major challenge for quantum chemistry. Following the previous studies [325–328], these states are denoted as  $8\pi(^1A_{1g})$ , a closed-shell singlet arising from the *sigma*-type  $b_{1g}$  HOMO being doubly occupied, having eight  $\pi$  electrons in its electron configuration,  $10\pi(^1A_{1g})$ , another closed-shell singlet where the  $\pi$ -type  $a_{2u}$  LUMO of the  $8\pi(^1A_{1g})$  reference is doubly occupied and the *sigma*-type  $b_{1g}$

HOMO is empty, having ten  $\pi$  electrons in its dominant electron configuration,  $9\pi(^3B_{2u})$ , the triplet state originating from the singly occupied HOMO and LUMO, having nine  $\pi$  electrons in its leading electron configuration, and  $9\pi(^1B_{2u})$ , an open-shell singlet counterpart of the  $9\pi(^3B_{2u})$  state, having nine  $\pi$  electrons in its dominant electron configuration (see Fig. 4.6).

Determining the singlet-triplet gap in  $C_4O_4$  along with the ordering of the four lowest-energy electronic states has proven to be very difficult for the various theoretical and experimental studies [325–331, 357, 358]. The initial DFT, CASSCF, CASPT2, and CCSD(T)-type calculations gave strongly varied results depending on the geometry employed and the method and basis set that were used [327, 328]. Given the challenge represented by the low-lying electronic states of  $C_4O_4$ , which require a robust and well-balanced description of dynamical and nondynamical correlation effects, this is not surprising. The DFT methodology, as implemented in the popular quantum chemistry codes, is known to have problems with electronic near degeneracies, CASSCF does not capture dynamical correlations, CASPT2 over-stabilizes singlet states with a strong biradical character relative to the corresponding single-reference closed-shell singlet and high-spin triplet states, and CCSD(T), which provides an accurate description of dynamical electron correlation effects for nondegenerate ground states of molecules near the equilibrium geometries, fails when applied to biradicals and other cases of electronic near degeneracies, where nondynamical correlation effects become more significant. The subsequent negative ion photoelectron (NIPE) spectroscopy [329] experiment determined the ground state to be the  $9\pi(^3B_{2u})$  state, being about  $6.27 \pm 0.5$  kJ/mol lower in energy than the lowest singlet state, in agreement with some of the theoretical results, but none of the experimental energy differences and even best theo-



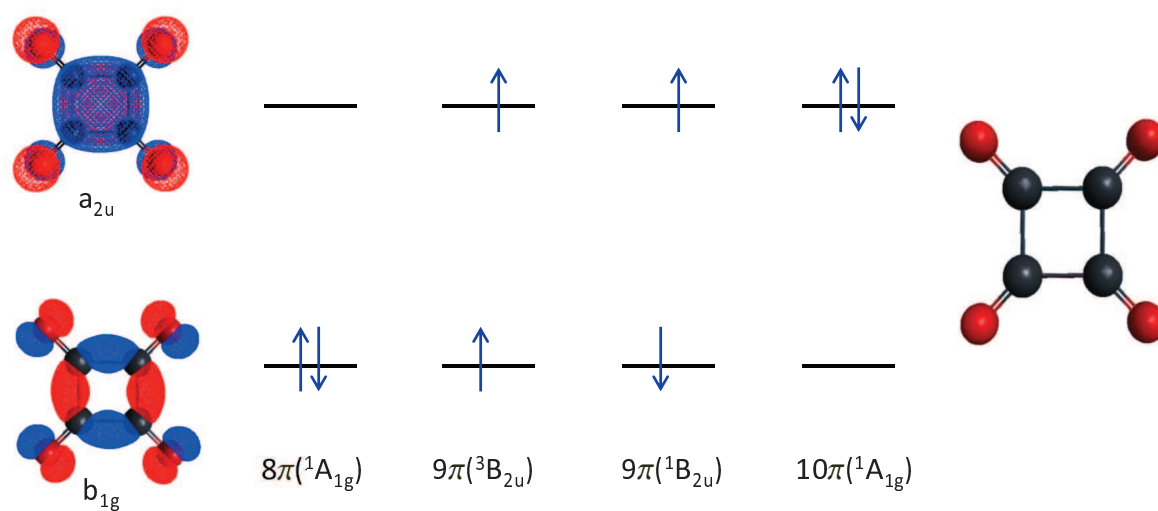


Figure 4.6: Electron distributions among the nearly degenerate HOMO and LUMO making up the four lowest-energy electronic states of  $C_4O_4$ .

retical predictions were in quantitative agreement. Furthermore, the original experimental interpretation [329] assigned a series of peaks seen in the NIPE spectrum of  $\text{C}_4\text{O}_4$  to the open-shell singlet  $9\pi(^1\text{B}_{2u})$  state [329], while in a follow-up paper, which included a detailed Franck-Condon analysis [330], the peaks were reassigned as a vibrational progression in the lowest-energy  $^1\text{A}_{1g}$  state, namely  $8\pi(^1\text{A}_{1g})$ , showing that even the experimental interpretation may be uncertain in some aspects.

More recently, to try to give a more definitive result for the singlet-triplet gap in  $\text{C}_4\text{O}_4$ , the high-level DEA- and DIP-EOMCC approaches, in which the EOM operators attaching two electrons to or removing two electrons from the corresponding closed-shell cores are truncated at  $3p-1h$  and  $3h-2p$  components, respectively, were applied to the  $8\pi(^1\text{A}_{1g})$  and  $9\pi(^3\text{B}_{2u})$  states in  $\text{C}_4\text{O}_4$  [331]. The authors of Ref. [331] found that even the DEA-EOMCC( $3p-1h$ ) and DIP-EOMCC( $3h-1p$ ) approximations were not sufficient to provide reliable results, with the  $8\pi(^1\text{A}_{1g})$  being predicted to be the ground state in some cases. This is not entirely surprising, though. For the DEA- and DIP-EOMCC methods to provide stable results, especially in situations as complex as that created by  $\text{C}_4\text{O}_4$ , one needs to include some form of  $4p-2h$  and  $4h-2p$  correlations, respectively, as has been shown for other challenging systems with low-lying nearly degenerate electronic states [177, 178]. This would explain difficulties in stabilizing the results of the DEA- and DIP-EOMCC calculations reported in Ref. [331], which did not include  $4p-2h$  and  $4h-2p$  excitations, but we would have to prove this via direct computations using, for example, the active-space DEA-EOMCC( $4p-2h$ ) and DIP-EOMCC( $4h-2p$ ) approaches developed in our group [177, 178]. Our present DEA-EOMCC( $4p-2h$ ) and DIP-EOMCC( $4h-2p$ ) codes have a pilot character and, as such, are

hard to apply to the larger-basis set calculations for  $\text{C}_4\text{O}_4$ , but this may be an interesting direction to pursue in future work.

All of this shows that  $\text{C}_4\text{O}_4$  is a remarkably challenging molecule for both theory and experiment. As far as theory is concerned, one family of CC methods, which has been shown to provide a balanced description of dynamical and nondynamical correlation effects, particularly in the case of biradicals and near-degenerate electronic states, is the previously discussed CR-CC hierarchy, including methods such as CR-CC(2,3), which we used in the earlier sections of this dissertation. It is relevant to mention in this context that the CR-CC(2,3) approach has already been used to determine the singlet-triplet gap [95] and, more recently, using its CR-EOMCC(2,3) extension, four lowest-energy states of methylene [192], achieving great success. Thus, as part of this dissertation, we examined the four lowest-energy states of  $\text{C}_4\text{O}_4$ , namely,  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^3\text{B}_{2u})$ ,  $9\pi(^1\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$ , investigating the role of both basis set and nuclear geometry to determine their ordering, by employing the CR-CC(2,3),A and CR-CC(2,3),D approaches that seek to minimize the difference between the exact, full CI, and CCSD energies by adding robust noniterative corrections due to triply excited clusters to the CCSD energy exploiting moments of CCSD equations discussed in Section 4.1. We then compare the CR-CC(2,3) results with multireference CI calculations using the popular MRCI(Q) method [359,360], the active-space CCSDt approach [62–67,361], which is known to precisely reproduce relative energies of its parent full CCSDT calculations, and the available experimental data [329].

We have already discussed the CR-CC(2,3) methodology in Section 4.1, but we have not said much about CCSDt. We do this now. The CCSDt method is based on the more general

active-space CC ideas developed in Refs. [62–64, 66, 69] for the ground state and [42–44] for excited states (for reviews of the active-space CC/EOMCC theories and their EA/IP, DEA/DIP, etc. extensions, and additional references, see Refs. [54, 157–159, 177, 178]). CCSDt allows one to obtain molecular electronic states that overall are of the full CCSDT quality at a small fraction of the computational cost involved in the CCSDT calculations. In the active-space CC/EOMCC methods, the spin-orbitals are divided into four disjoint groups, namely, (i) core or inactive occupied spin-orbitals, labeled by bold lower-case letters **i**, **j**, ... , (ii) active occupied spin-orbitals, labeled by upper-case bold letters **I**, **J**, ... , (iii) active unoccupied spin-orbitals, labeled by upper-case bold letters **A**, **B**, ... , and (iv) virtual or inactive unoccupied spin-orbitals, labeled by bold lower-case letters **a**, **b**, ... (see Fig. 4.7). We continue to designate the occupied and unoccupied orbitals in the reference determinant  $|\Phi\rangle$  by the italic lower-case characters  $i, j, \dots$  and  $a, b, \dots$ , respectively, if the active/inactive character of the spin-orbitals is not specified. We then use the above decomposition of spin-orbitals into active and inactive subsets to capture the most important higher-than-double excitations, such as triples (CCSDt) or triples and quadruples (CCSDtq), reducing the overall cost of high-level CC computations with those excitations, such as CCSDT or CCSDTQ. In the specific case of CCSDt, which is an approximation to full CCSDT, the one- and two-body cluster operators,  $T_1$  and  $T_2$ , are defined through exactly the same set of amplitudes as used in the standard CCSD calculations, discussed in Section 4.1, with their labels running over all spin-orbitals. In other words,  $T_1$  and  $T_2$  are treated fully. A different approach, however, is adopted when it comes to  $T_3$  clusters (triples), which in CCSDt are replaced by their  $t_3$  (“little t”) counterparts defined through a small subset

of all triply excited cluster amplitudes that carry at least one active occupied and at least one active unoccupied spin-orbital indices, to obtain the following definition of the cluster operator  $T$ :

$$T^{(\text{CCSDt})} = T_1 + T_2 + t_3 = \sum_{i,a} t_a^i a^a a_i + \sum_{i>j, a>b} t_{ab}^{ij} a^a a^b a_j a_i + \sum_{\mathbf{I}>j>k, a>b>\mathbf{C}} t_{ab\mathbf{C}}^{\mathbf{I}jk} a^a a^b a^{\mathbf{C}} a_k a_j a_{\mathbf{I}}, \quad (4.30)$$

where  $a^p$  ( $a_p$ ) are the same creation (annihilation) operators as defined previously associated with the spin-orbital basis set  $\{|p\rangle\}$ . The ground-state wave function is then written as  $|\Psi_0^{(\text{CCSDt})}\rangle = e^{T^{(\text{CCSDt})}}|\Phi\rangle$  and we solve a set of nonlinear equations similar to CCSD, namely,  $\langle\Phi_i^a|\bar{H}^{(\text{CCSDt})}|\Phi\rangle = 0$ ,  $\langle\Phi_{ij}^{ab}|\bar{H}^{(\text{CCSDt})}|\Phi\rangle = 0$ , and  $\langle\Phi_{\mathbf{I}jk}^{ab\mathbf{C}}|\bar{H}^{(\text{CCSDt})}|\Phi\rangle = 0$ , where  $\bar{H}^{(\text{CCSDt})}$  is the similarity transformed Hamiltonian of CC theory written for the cluster operator  $T^{(\text{CCSDt})}$  and  $|\Phi_i^a\rangle$ ,  $|\Phi_{ij}^{ab}\rangle$ , and  $|\Phi_{\mathbf{I}jk}^{ab\mathbf{C}}\rangle$  are the singly, doubly, and triply excited determinants, respectively, corresponding to the cluster amplitudes defining  $T^{(\text{CCSDt})}$ . As one can see, the CCSDt equations form a subset of the full CCSDT equations and become equivalent to the latter equations when all spin orbitals are chosen as active.

The key idea of the CCSDt calculations is that one does not have to use the entire set of triply excited amplitudes  $t_{abc}^{ijk}$  and determinants  $|\Phi_{ijk}^{abc}\rangle$  to obtain the results of the full CCSDT quality, which the definition of  $t_3$  operator shown in Eq. (4.30) guarantees (as shown in numerous applications). Because of the considerable reduction in the number of triples used by CCSDt, the most expensive steps of CCSDt scale as  $N_o N_u n_o^2 n_u^4$ , where  $n_o$  ( $N_o$ ) and  $n_u$  ( $N_u$ ) are the numbers of all (active) occupied and unoccupied spin-orbitals used in the post-SCF calculations. Since one typically has that  $N_o < n_o$  and  $N_u \ll n_u$ , the CPU costs of CCSDt calculations are the costs of CCSD times a small prefactor, which

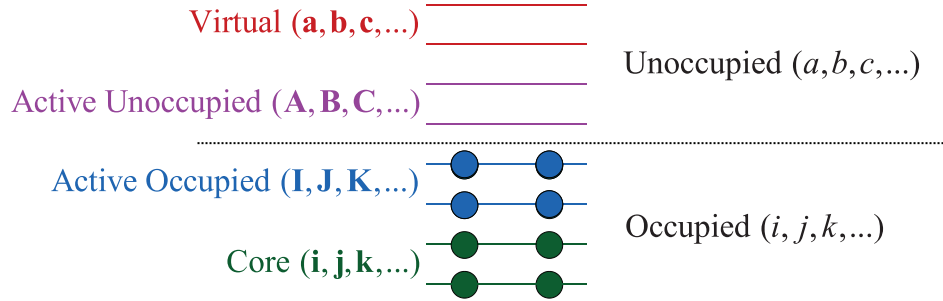


Figure 4.7: Orbital classification used in the active-space CC approaches, such as CCSDt. Full circles represent core and active electrons of the reference determinant  $|\Phi\rangle$ , which is the formal reference defining the Fermi vacuum for the active-space CC calculations.

is a lot less than the  $n_o^3 n_u^5$  steps of CCSDT. At the same time, by using a small subset of triples, the CCSDt calculations reduce the  $n_o^3 n_u^3$  storage requirements of CCSDT to a much smaller  $\sim N_o N_u n_o^2 n_u^2$  level. In the rest of our discussion of the CCSDt results, the CCSDt calculations using  $N_o$  active occupied and  $N_u$  active unoccupied orbitals are designated as CCSDt( $N_o, N_u$ ). Clearly, CCSDt( $N_o, N_u$ ) becomes CCSDT if  $N_o = n_o$  and  $N_u = n_u$ .

### 4.4.2 Computational details

Our goal was to determine the adiabatic excitation energies and ordering of the four low-lying electronic states of  $\text{C}_4\text{O}_4$ , including  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$ , using high-level *ab initio* methods based on CC theory, especially CR-CC(2,3) and CCSDt, augmented by CASSCF-based MRCI(Q) calculations, and examining the role of the basis set and methods exploited in geometry optimizations. The  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ , and  $10\pi(^1A_{1g})$  states were calculated with the ground-state CR-CC(2,3) and CCSDt methods, since it was easy to determine suitable reference determinants for each of these three states. Thus, in the case of the  $8\pi(^1A_{1g})$  state, we used the RHF reference determinant corresponding to the orbital occupancy shown in the left-most configuration of Fig. 4.6. In the case of the  $9\pi(^3B_{2u})$  state, we used the ROHF orbitals corresponding to the high-spin triplet configuration, shown in Fig. 4.6 as well. For the  $10\pi(^1A_{1g})$  state dominated by the determinant shown in Fig. 4.6 as the right-most configuration, we used the RHF reference corresponding to the orbital occupancy defining this right-most configuration, obtained by interchanging the HOMO and LUMO in the  $8\pi(^1A_{1g})$  reference. The above choices of reference determinants for the CR-CC(2,3) and CCSDt calculations (and their CCSD counterparts) for the  $8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g})$  states, where each determinant is optimized for the electronic state it approximates, are justified by the early observation [327] that the closed-shell electron configurations dominating the  $8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g})$  states do not mix. In the case of the  $9\pi(^3B_{2u})$  state, we rely on its obviously single-reference character, where the high-spin ROHF determinant is the optimum spin- and symmetry-adapted reference.

The only state that required a different approach was the open-shell singlet  $9\pi(^1B_{2u})$

state. One could, of course, try the unrestricted CC (UCC) calculations using one of the two leading determinants in the  $9\pi(^1B_{2u})$  wave function, such as that shown in the third configuration from the left in Fig. 4.6, as a UHF reference. Unfortunately, the resulting UCC solutions would lead to complete symmetry breakdown and significant spin contamination that would be hard to remove, raising several questions about the meaning of such results. Another approach that one can take, which avoids problems associated with symmetry breaking and spin contamination, is to treat the  $9\pi(^1B_{2u})$  state as an excited state out of the closed-shell  $8\pi(^1A_{1g})$  state. A state like this can be obtained by diagonalizing the similarity transformed Hamiltonian obtained in the closed-shell CC calculations for the  $8\pi(^1A_{1g})$  state, as in the EOMCC considerations. The  $9\pi(^1B_{2u})$  state is an excited state relative to the  $8\pi(^1A_{1g})$  state dominated by a one-electron transition from the  $b_{1g}$  HOMO to the  $a_{2u}$  LUMO. The basic EOMCCSD approach usually provides a reasonable description of excited states dominated by one-electron transitions. Thus, we decided to determine the  $9\pi(^1B_{2u})$  state as an excitation from the closed-shell  $8\pi(^1A_{1g})$  state obtained using EOMCCSD. This has allowed us to treat the  $9\pi(^1B_{2u})$  state in a proper spin- and symmetry-adapted manner, i.e., without worries about symmetry breaking and spin contamination and without *ad hoc*, at best qualitative, arguments based on examining the CCSD(T) results and exchange integrals presented in Ref. [327]. In analogy to the remaining three states, we should, at least in principle, correct the EOMCCSD results for the  $8\pi(^1A_{1g}) \rightarrow 9\pi(^1B_{2u})$  excitation for the effects due to triples, either through the  $(\delta)$ -CR-EOMCC(2,3), EOMCCSD( $\tilde{T}$ ), EOMCCSD(2)<sub>T</sub>, and similar noniterative corrections or by running the iterative EOMCCSDt calculations, but our preliminary tests have demonstrated a large variation among the results in this



case, beyond acceptable levels, and unexpected difficulties in running EOMCCSDt with the software we used, so we will return to this issue in a future work. The  $9\pi(^1B_{2u})$  state is an almost pure one-electron transition of the same spin symmetry as the  $8\pi(^1A_{1g})$  state, so we do not expect the effects due to triples to significantly alter our main conclusions based on the EOMCCSD calculations for the  $8\pi(^1A_{1g}) \rightarrow 9\pi(^1B_{2u})$  excitation (in fact, as shown below, they agree quite well with MRCI(Q)), but additional EOMCC studies regarding this excitation may be needed in the future.

Once we decided on the appropriate ways of treating the  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$  states of  $C_4O_4$ , including reference determinants employed in the CC calculations, we moved to a series of single-point energy calculations at the UB3LYP, CASSCF(2,2), and CASSCF(16,16) optimized geometries, obtained with the 6-31G(d) basis set, taken from Ref. [327], using the CCSD (EOMCCSD for the  $9\pi(^1B_{2u})$  state), CR-CC(2,3),A, and CR-CC(2,3),D methods with the 6-31G(d) and aug-cc-pV $x$ Z ( $x = D, T$ ) bases [186, 362], the active-space CCSDt(8,8) and CCSDt(16,16) approaches employing the 6-31G(d) and aug-cc-pVDZ basis sets, and the state-averaged CASSCF(16,16) and MRCI(Q)(14,13)//CASSCF(16,16) methods using the aug-cc-pVDZ basis, where the equal state-averaging of the  $8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g})$  states was employed (see below for discussion of the active spaces chosen). The CCSDt and MRCI(Q) calculations employing the aug-cc-pVTZ basis set were not performed due to their prohibitive computational costs, when an adequate active space to capture the relevant excitations is used (even if run in parallel). All of the calculated energies of the  $8\pi(^1A_{1g})$  state and the adiabatic excitation energies for the remaining three low-lying electronic states of  $C_4O_4$  are presented in Table 4.9.

Table 4.9: Energies of the four lowest-energy electronic states of  $C_4O_4$  at the UB3LYP/6-31G(d), CASSCF(2,2)/6-31G(d), and CASSCF(16,16)/6-31G(d) geometries obtained in Ref. [327].<sup>a</sup>

| Method                        | Basis set   | UB3LYP/6-31G(d) geometry |                  |                  |                   | CASSCF(2,2)/6-31G(d) geometry |                  |                  |                   | CASSCF(16,16)/6-31G(d) geometry |                  |                  |                   |
|-------------------------------|-------------|--------------------------|------------------|------------------|-------------------|-------------------------------|------------------|------------------|-------------------|---------------------------------|------------------|------------------|-------------------|
|                               |             | $8\pi(^1A_{1g})$         | $9\pi(^3B_{2u})$ | $9\pi(^1B_{2u})$ | $10\pi(^1A_{1g})$ | $8\pi(^1A_{1g})$              | $9\pi(^3B_{2u})$ | $9\pi(^1B_{2u})$ | $10\pi(^1A_{1g})$ | $8\pi(^1A_{1g})$                | $9\pi(^3B_{2u})$ | $9\pi(^1B_{2u})$ | $10\pi(^1A_{1g})$ |
| CCSD <sup>b</sup>             | 6-31G(d)    | 1.9968814                | 30.093           | 35.648           | 116.803           | 1.9888584                     | 33.276           | 41.368           | 121.106           | 1.9966172                       | 40.155           | 47.928           | 124.482           |
|                               | aug-cc-pVDZ | 2.1488333                | 17.222           | 21.588           | 94.604            | 2.1412594                     | 16.780           | 23.202           | 93.794            | 2.1484649                       | 22.166           | 27.931           | 99.343            |
|                               | aug-cc-pVTZ | 2.4915799                | 21.185           | 23.760           | 100.740           | 2.4893788                     | 19.594           | 24.446           | 99.209            | 2.4919414                       | 20.265           | 25.106           | 97.480            |
| CR-CC(2,3),A                  | 6-31G(d)    | 2.0410444                | 4.130            | -                | 51.413            | 2.0306655                     | 10.010           | -                | 63.325            | 2.0404453                       | 18.563           | -                | 66.881            |
|                               | aug-cc-pVDZ | 2.2008911                | -8.725           | -                | 31.261            | 2.1907084                     | -6.356           | -                | 37.774            | 2.2001588                       | 0.975            | -                | 43.969            |
|                               | aug-cc-pVTZ | 2.5638980                | -6.752           | -                | 33.767            | 2.5587924                     | -5.354           | -                | 39.918            | 2.5638591                       | -2.368           | -                | 39.068            |
| CR-CC(2,3),D                  | 6-31G(d)    | 2.0459569                | 0.870            | -                | 43.125            | 2.0353747                     | 6.866            | -                | 55.431            | 2.0453289                       | 15.612           | -                | 59.049            |
|                               | aug-cc-pVDZ | 2.2054282                | -12.158          | -                | 23.845            | 2.1948632                     | -10.091          | -                | 30.558            | 2.2045827                       | -2.303           | -                | 36.955            |
|                               | aug-cc-pVTZ | 2.5678525                | -10.633          | -                | 27.502            | 2.5628428                     | -8.301           | -                | 34.628            | 2.5678184                       | -5.895           | -                | 33.220            |
| CCSDt(8,8)                    | 6-31G(d)    | 2.0328825                | 2.923            | -                | 49.895            | 2.0222028                     | 8.755            | -                | 60.630            | 2.0322461                       | 17.272           | -                | 65.323            |
|                               | aug-cc-pVDZ | 2.1735390                | -11.417          | -                | 25.207            | 2.1612823                     | -14.305          | -                | 24.090            | 2.1728430                       | -0.619           | -                | 36.764            |
| CCSDt(16,16)                  | 6-31G(d)    | 2.0414828                | 2.073            | -                | 46.564            | 2.0308715                     | 8.194            | -                | 58.165            | 2.0408498                       | 16.906           | -                | 62.183            |
|                               | aug-cc-pVDZ | 2.1771695                | -9.486           | -                | 25.173            | 2.1677944                     | -6.124           | -                | 32.558            | 2.1766114                       | 0.932            | -                | 37.181            |
| CASSCF(16,16) <sup>c</sup>    | aug-cc-pVDZ | 1.0605689                | 59.132           | 91.116           | 141.514           | 1.0602076                     | 59.790           | 92.456           | 149.577           | 1.0613717                       | 60.195           | 93.782           | 141.161           |
| MRCI(Q)(14,13) <sup>c,d</sup> | aug-cc-pVDZ | 1.1876699                | -7.808           | 29.853           | 43.446            | 1.1893300                     | 0.358            | 37.573           | 62.983            | 1.1888253                       | -2.252           | 36.927           | 46.627            |

<sup>a</sup>The  $8\pi(^1A_{1g})$  total electronic energies (in hartree) are reported as  $(-450 - E)$ , while the energies of the remaining three states are adiabatic excitation energies, in kJ/mol, relative to the  $8\pi(^1A_{1g})$  state.

<sup>b</sup>The  $9\pi(^1B_{2u})$  adiabatic excitation energies were calculated as the differences between the  $9\pi(^1B_{2u})$  EOMCCSD and  $8\pi(^1A_{1g})$  CCSD energies.

<sup>c</sup>State averaging of the  $8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g})$  states was used.

<sup>d</sup>MRCI(Q) single-point energies obtained using CASSCF(16,16)/aug-cc-pVDZ optimized orbitals (see text for details).

Let us briefly comment on the active spaces used in the CCSDt, CASSCF, and MRCI(Q) calculations. The CCSDt(8,8) active space was chosen such that it consisted of the same types of active orbitals as those used in the CASSCF(16,16) computations reported in Ref. [327], i.e., the 8 highest-energy occupied orbitals and the 8 lowest-energy unoccupied orbitals of each of the three states of  $\text{C}_4\text{O}_4$  determined with CCSDt (the only difference between the CCSDt and CASSCF-based calculations was that CCSDt used RHF orbitals, not the multiconfigurational ones). This was also the smallest meaningful active space one could use to obtain reliable results for the low-lying electronic states of  $\text{C}_4\text{O}_4$  (we tested this by considering smaller active spaces). To determine the effect of the size of the active space on the convergence of our CCSDt results toward the full CCSDT solutions, we also performed the considerably large CCSDt(16,16) calculations, employing the 16 highest-energy occupied orbitals and the 16 lowest-energy unoccupied orbitals in the active space, which would correspond to multireference calculations based on CASSCF(32,32) references, i.e., the calculations that are not feasible at this time due to the factorial scaling of the dimensionality of CASSCF wave functions with the numbers of active orbitals and correlated active electrons (this illustrates one of the key advantages of the active-space CC methods over the CASSCF-based multireference approaches; the active-space CC methods have a low-order polynomial scaling of the CPU operational count and dimensionalities of wave function spaces used to set them up with the numbers of active occupied and unoccupied orbitals). The active space for the MRCI(Q)(14,13) calculations was constructed from the 7 highest-energy occupied orbitals and 6 lowest-energy unoccupied orbitals of each of the four electronic states of interest in this study. We chose the somewhat smaller (14,13) active space

in this case, and not the (16,16) one used in the preceding CASSCF orbital optimizations, due to this being the largest computationally feasible active space that we could afford at the MRCI(Q) level that did not fall across degenerate orbital pairs for all four electronic states of interest. The orbital composition of the active space for each electronic state can be found in Table 4.10.

After performing the above single-point CC and MRCI calculations using UB3LYP and CASSCF geometries, we performed a series of geometry optimizations for the four low-lying electronic states of  $\text{C}_4\text{O}_4$  employing the CCSD, CR-CC(2,3),A, and CR-CC(2,3),D approaches for the  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ , and  $10\pi(^1A_{1g})$  states and EOMCCSD for the  $9\pi(^1B_{2u})$  state, using the 6-31G(d) and aug-cc-pVDZ basis sets. These high-level *ab initio* optimizations were carried out using the parallel coarse-grain finite-difference model available in the CIOpt program suite [235,363], which we interfaced with our relevant CC/EOMCC routines available in the GAMESS package. The CC-level optimized geometries for the four states of  $\text{C}_4\text{O}_4$  examined in this work are summarized in Table 4.11, along with the geometries optimized in the previous study [327] shown for comparison.

Table 4.10: Orbital structure of CASSCF(16,16) and MRCI(Q)(14,13) active spaces.

| Method            | Active Orbitals and Occupancy in the Leading Determinants                        |                                                                                  |
|-------------------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
|                   | CASSCF(16,16)                                                                    |                                                                                  |
| $8\pi(^1A_{1g})$  | $(1a_{2u}^2)(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^2)$ | $(2a_{2u}^0)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)(3e_g^0)(4e_g^0)$ |
| $9\pi(^3B_{2u})$  | $(1a_{2u}^2)(1e_g^2)(2e_g^2)(1b_{1u}^2)(1a_{2g}^2)(11e_u^2)(12e_u^2)(2b_{1g}^1)$ | $(2a_{2u}^1)(3e_g^0)(4e_g^0)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)$ |
| $9\pi(^1B_{2u})$  | $(1a_{2u}^2)(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^1)$ | $(2a_{2u}^1)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)(3e_g^0)(4e_g^0)$ |
| $10\pi(^1A_{1g})$ | $(1a_{2u}^2)(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^0)$ | $(2a_{2u}^2)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)(3e_g^0)(4e_g^0)$ |
|                   | MRCI(Q)(14,13)                                                                   |                                                                                  |
| $8\pi(^1A_{1g})$  | $(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^2)$            | $(2a_{2u}^0)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)$                 |
| $9\pi(^3B_{2u})$  | $(1e_g^2)(2e_g^2)(1b_{1u}^2)(1a_{2g}^2)(11e_u^2)(12e_u^2)(2b_{1g}^1)$            | $(2a_{2u}^1)(3e_g^0)(4e_g^0)(6a_{1g}^0)(13e_u^0)(14e_u^0)$                       |
| $9\pi(^1B_{2u})$  | $(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^1)$            | $(2a_{2u}^1)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)$                 |
| $10\pi(^1A_{1g})$ | $(1a_{2g}^2)(1e_g^2)(2e_g^2)(11e_u^2)(12e_u^2)(1b_{1u}^2)(2b_{1g}^0)$            | $(2a_{2u}^2)(6a_{1g}^0)(13e_u^0)(14e_u^0)(3a_{2u}^0)(5b_{2g}^0)$                 |

Table 4.11: Bond lengths of optimized  $C_4O_4$   $D_{4h}$  geometries (in Å).<sup>a</sup>

| Method                     | Basis       | $8\pi(^1A_{1g})$ | $9\pi(^3B_{2u})$ | $9\pi(^1B_{2u})$ | $10\pi(^1A_{1g})$ |
|----------------------------|-------------|------------------|------------------|------------------|-------------------|
| UB3LYP <sup>b</sup>        | 6-31G(d)    | (1.567,1.200)    | (1.553,1.196)    | (1.552,1.196)    | (1.546,1.190)     |
| CASSCF(2,2) <sup>b</sup>   | 6-31G(d)    | (1.551,1.173)    | (1.562,1.160)    | (1.562,1.160)    | (1.571,1.151)     |
| CASSCF(16,16) <sup>b</sup> | 6-31G(d)    | (1.562,1.197)    | (1.576,1.172)    | (1.581,1.172)    | (1.583,1.166)     |
| CCSD <sup>c</sup>          | 6-31G(d)    | (1.560,1.206)    | (1.556,1.196)    | (1.556,1.198)    | (1.554,1.187)     |
|                            | aug-cc-pVDZ | (1.574,1.203)    | (1.571,1.192)    | (1.570,1.194)    | (1.568,1.184)     |
| CR-CC(2,3),A               | 6-31G(d)    | (1.565,1.212)    | (1.558,1.203)    | -                | (1.551,1.198)     |
|                            | aug-cc-pVDZ | (1.578,1.209)    | (1.574,1.200)    | -                | (1.566,1.195)     |
| CR-CC(2,3),D               | 6-31G(d)    | (1.565,1.212)    | (1.558,1.204)    | -                | (1.551,1.198)     |
|                            | aug-cc-pVDZ | (1.579,1.210)    | (1.573,1.201)    | -                | (1.565,1.196)     |

<sup>a</sup>Geometries are reported as ( $R_{C-C}, R_{C-O}$ ) pairs.

<sup>b</sup>Taken from Ref. [327].

<sup>c</sup>Geometry of the  $9\pi(^1B_{2u})$  state was optimized at the EOMCCSD level of theory.

After performing the above geometry optimizations at the CCSD/EOMCCSD and CR-CC(2,3),X ( $X = A, D$ ) levels, we moved to the final round of single-point CC/aug-cc-pVTZ energy calculations using the corresponding CC/aug-cc-pVDZ equilibrium geometries. This was done to assess the stability of our CCSD/EOMCCSD and CR-CC(2,3) optimized results using the aug-cc-pVDZ basis with respect to the basis set. We also computed approximate zero-point vibrational energy (ZPE) corrections at the UB3LYP/aug-cc-pVTZ level to see if there was an appreciable effect on the state orderings from vibrations [364]. The resulting adiabatic excitation energies for the  $9\pi(^3B_{2u})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$  states, with respect to the  $8\pi(^1A_{1g})$  state, from our various CC/aug-cc-pVDZ nuclear geometry optimizations and subsequent single-point CC/aug-cc-pVTZ energy calculations are presented in Table 4.12.

All CCSD, EOMCCSD, CR-CC(2,3) and UB3LYP calculations were carried out using

the GAMESS package, where our CC/EOMCC codes reside. The CCSDt computations were performed with the NWChem [322] suite. The CASSCF and MRCI(Q) results were obtained using the Molpro [365] program. In all of the post-SCF calculations, the core electrons were not correlated and for all calculations performed in this work the spherical components of  $d$  and  $f$  basis functions were employed throughout.

### 4.4.3 Results and discussion

The results of our initial single-point CC energy calculations, using the UB3LYP, CASSCF(2,2), and CASSCF(16,16) optimized geometries employing the 6-31G(d) basis set, are shown in Table 4.9. When we look at our best CR-CC(2,3), CCSDt, and MRCI(Q) calculations, it is clear that one needs to use a basis set of at least the aug-cc-pVDZ or aug-cc-pVTZ quality to obtain a reasonably converged and correct description of state ordering. Indeed, if we use the small 6-31G(d) basis set in the CR-CC(2,3), CCSDt, and MRCI(Q) calculations, the  $8\pi(^1A_{1g})$  state becomes the ground state, independent of the method used to optimize the geometries, contradicting the experimental findings [329, 330]. The situation changes when we use the larger aug-cc-pVDZ and aug-cc-pVTZ basis sets in the CR-CC(2,3), CCSDt, and MRCI(Q) calculations, which clearly favor the  $9\pi(^3B_{2u})$  state as the ground state, placing it about 6 – 11 kJ/mol below the  $8\pi(^1A_{1g})$  state in the majority of our best CR-CC(2,3),D/aug-cc-pVTZ calculations (the CR-CC(2,3),A/aug-cc-pVTZ calculations give a 2 – 7 kJ/mol range below the  $8\pi(^1A_{1g})$  state), 6 – 9 kJ/mol below the  $8\pi(^1A_{1g})$  state in the two of the three single-point CCSDt(16,16)/aug-cc-pVDZ results, and 2 – 8 kJ/mol below the  $8\pi(^1A_{1g})$  state in the two of the three MRCI(Q)(14,13)/aug-cc-pVDZ calculations. There

are a couple of cases, such as the CCSDt(16,16)/aug-cc-pVDZ//CASSCF(16,16)/6-31G(d) and MRCI(Q)(13,14)/aug-cc-pVDZ//CASSCF(2,2)/6-31G(d) results, where the  $8\pi(^1A_{1g})$  and  $9\pi(^3B_{2u})$  state become virtually degenerate, with the latter state being slightly higher in energy, but we must keep in mind that these are single-point calculations at the geometries obtained with the unrelated methods. Even with these few exceptions, based on the results in Table 4.9, it is reasonable to conclude that the  $9\pi(^3B_{2u})$  state is the ground state, with the  $8\pi(^1A_{1g})$  state being the first excited state. We must, however, rely on the higher-level methods, such as CR-CC(2,3), CCSDt, or MRCI(Q), and basis sets of the aug-cc-pVDZ or, preferably, aug-cc-pVTZ quality to obtain a correct description. The smaller basis sets, such as 6-31G(d), and lower-order CC methods, such as CCSD, which favor the  $8\pi(^1A_{1g})$  ground state, are clearly inadequate in this regard.

There are other useful observations based on the results collected in Table 4.9. In the vast majority of calculations reported in Table 4.9, the change of the active space from  $(N_o, N_u) = (8, 8)$  to  $(N_o, N_u) = (16, 16)$  has no effect on the results of the CCSDt calculations, which typically vary by 1 – 2 kJ/mol, when the  $9\pi(^3B_{2u}) - 8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g}) - 8\pi(^1A_{1g})$  energy differences are considered, showing that our best active-space CCSDt(16,16)/aug-cc-pVDZ calculations are likely to be converged to within 1 – 2 kJ/mol for these two energy differences (i.e., being very close to the corresponding CCSDT results, which we could not obtain due to prohibitive costs of full CCSDT calculations for  $C_4O_4$ ). This is a very valuable piece of information, since there is an equally high degree of consistency between the high-level CCSDt(16,16)/aug-cc-pVDZ and CR-CC(2,3),X/aug-cc-pVDZ or CR-CC(2,3),X/aug-cc-pVTZ ( $X = A, D$ ) results for the  $9\pi(^3B_{2u}) - 8\pi(^1A_{1g})$  and  $10\pi(^1A_{1g}) - 8\pi(^1A_{1g})$  energy



differences, allowing us to further increase our trust in the CR-CC(2,3) results obtained with the aug-cc-pVDZ and aug-cc-pVTZ basis sets.

Given the overall picture that emerges from Table 4.9, the  $9\pi(^3B_{2u})$  state is expected to be the ground state, the  $8\pi(^1A_{1g})$  state is the first excited state, and the  $10\pi(^1A_{1g})$  state is the highest in energy among the four states of  $C_4O_4$  examined in this work. All of the highest-level methods considered in this study, i.e., CR-CC(2,3), CCSDt, and MRCI(Q), place the  $10\pi(^1A_{1g})$  state at about 30 – 40 kJ/mol above the  $8\pi(^1A_{1g})$  state when the aug-cc-pVDZ and aug-cc-pVTZ basis sets are employed. The only state in Table 4.9, which is treated differently than the remaining  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ , and  $10\pi(^1A_{1g})$  states, is the open-shell singlet  $9\pi(^1B_{2u})$  state. To avoid problems with spin-symmetry breaking, which the UCC computations (including CR-CC(2,3) and CCSDt levels) would cause, we treated it as an excitation from  $8\pi(^1A_{1g})$  described by EOMCCSD. Although clearly not definitive due to the neglect of triples in this case, the EOMCCSD results for the  $9\pi(^1B_{2u})$  state obtained with the aug-cc-pVDZ and aug-cc-pVTZ basis sets shown in Table 4.9 agree quite well with the high-level (also, spin- and symmetry-adapted) MRCI(Q)(14,13)/aug-cc-pVDZ calculations, placing it about 20 – 30 kJ/mol above the  $8\pi(^1A_{1g})$  state, always above the  $9\pi(^3B_{2u})$  and below the  $10\pi(^1A_{1g})$  states. Thus, we can treat an EOMCCSD description of the  $9\pi(^1B_{2u}) - 8\pi(^1A_{1g})$  energy difference as quite trustworthy, when the aug-cc-pVDZ and aug-cc-pVTZ basis sets are employed. Based on the above discussion, we can, at least tentatively at this point, conclude that the ordering of the four low-lying electronic states of  $C_4O_4$  is  $9\pi(^3B_{2u}) < 8\pi(^1A_{1g}) < 9\pi(^1B_{2u}) < 10\pi(^1A_{1g})$ .

The discrepancies between the UB3LYP, CASSCF(2,2), and CASSCF(16,16) geometries

Table 4.12: Energies of the low-lying electronic states of C<sub>4</sub>O<sub>4</sub> at the geometries obtained with CC methods.<sup>a</sup>

| Method                      | Basis                    | 8 $\pi(^1A_{1g})$ | 9 $\pi(^3B_{2u})$ | 9 $\pi(^1B_{2u})$           | 10 $\pi(^1A_{1g})$ |
|-----------------------------|--------------------------|-------------------|-------------------|-----------------------------|--------------------|
| CCSD <sup>b</sup>           | 6-31G(d)                 | -451.9971417      | 30.741            | 36.214                      | 117.085            |
|                             | aug-cc-pVDZ              | -452.1489716      | 16.193            | 20.613                      | 92.583             |
|                             | aug-cc-pVTZ <sup>c</sup> | -452.4909723      | 18.858            | 22.410                      | 97.235             |
| ZPE corrected <sup>d</sup>  |                          |                   | 18.804            | 24.399                      | 96.619             |
| MRCI(Q)(14,13) <sup>e</sup> | aug-cc-pVDZ              | -451.1860936      | -7.883            | 30.590                      | 45.452             |
| CR-CC(2,3),A                | 6-31G(d)                 | -452.0419014      | 5.350             | -                           | 52.420             |
|                             | aug-cc-pVDZ              | -452.2017392      | -11.929           | -                           | 31.071             |
|                             | aug-cc-pVTZ <sup>c</sup> | -452.5628388      | -7.488            | -                           | 33.486             |
| ZPE corrected <sup>d</sup>  |                          |                   | -7.542            | -                           | 32.870             |
| MRCI(Q)(14,13) <sup>e</sup> | aug-cc-pVDZ              | -451.1833388      | -8.000            | -                           | 43.374             |
| CR-CC(2,3),D                | 6-31G(d)                 | -452.0469052      | 2.190             | -                           | 44.143             |
|                             | aug-cc-pVDZ              | -452.2064023      | -11.931           | -                           | 23.840             |
|                             | aug-cc-pVTZ <sup>c</sup> | -452.5668383      | -11.087           | -                           | 27.359             |
| ZPE corrected <sup>d</sup>  |                          |                   | -11.141           | -                           | 26.743             |
| MRCI(Q)(14,13) <sup>e</sup> | aug-cc-pVDZ              | -451.1829964      | -9.202            | -                           | 42.583             |
| Experiment <sup>f</sup>     |                          |                   | -6.27 $\pm$ 0.5   | 7.24 $\pm$ 0.5 <sup>g</sup> | -                  |

<sup>a</sup>The 8 $\pi(^1A_{1g})$  total energies are in hartree, while the energies of the remaining three states are adiabatic excitation energies, in kJ/mol, relative to the 8 $\pi(^1A_{1g})$  state.

<sup>b</sup>The 9 $\pi(^1B_{2u})$  adiabatic excitation energies are the differences between the 9 $\pi(^1B_{2u})$  EOMCCSD and 8 $\pi(^1A_{1g})$  CCSD energies.

<sup>c</sup>Single-point energies at the corresponding CC/aug-cc-pVDZ optimized geometries.

<sup>d</sup>ZPE corrections were computed at the UB3LYP/aug-cc-pVTZ level.

<sup>e</sup>Single-point energies using CASSCF(16,16)/aug-cc-pVDZ optimized orbitals computed at the corresponding CC/aug-cc-pVDZ optimized geometries. The equally weighted state averaging of the 8 $\pi(^1A_{1g})$  and 10 $\pi(^1A_{1g})$  states was employed.

<sup>f</sup>Taken from Ref. [329].

<sup>g</sup>Originally assigned as the 8 $\pi(^1A_{1g})$  and 9 $\pi(^1B_{2u})$  energy difference in Ref. [329]. Later, in Ref. [330], it was shown to be due to a vibrational progression of the 8 $\pi(^1A_{1g})$  state.

shown in Table 4.11 and the variation among the results of the single-point CR-CC(2,3), CCSDt, and MRCI(Q) calculations at the UB3LYP, CASSCF(2,2), and CASSCF(16,16) geometries, reported in Table 4.9, demonstrate that geometry relaxation may have a non-negligible effect on the calculated energy differences between the low-lying electronic states of  $\text{C}_4\text{O}_4$  examined in this study. Thus, to be able to be more definitive about our remarks concerning the relative order of the three lowest-energy singlet states of  $\text{C}_4\text{O}_4$  and the gap between the triplet  $9\pi(^3\text{B}_{2u})$  and singlet  $8\pi(^1\text{A}_{1g})$  states, we carried out the additional geometry optimizations at the CCSD, CR-CC(2,3),A, and CR-CC(2,3),D levels for the  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^3\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$  states and the EOMCCSD level for the  $9\pi(^1\text{B}_{2u})$  state using the 6-31G(d) and aug-cc-pVDZ basis sets. Given the excellent agreement between the CR-CC(2,3) and CCSDt results for the  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^3\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$  states, where the latter results are expected to be close to their full CCSDT counterparts, and the overall good agreement between the EOMCCSD and MRCI(Q) calculations for the  $9\pi(^1\text{B}_{2u})$  state, the CR-CC(2,3) and EOMCCSD optimizations performed in this work should provide us with useful additional insights regarding the relative ordering of the  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^3\text{B}_{2u})$ ,  $9\pi(^1\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$  states of  $\text{C}_4\text{O}_4$ . The results of the above CC-level geometry optimizations, along with the UB3LYP, CASSCF(2,2), and CASSCF(16,16) geometries from the previous study [327], are shown in Table 4.11.

As demonstrated in Table 4.11, our best CR-CC(2,3)/aug-cc-pVDZ geometries differ from the previous reported UB3LYP/6-31G(d) and CASSCF/6-31G(d) results rather significantly, demonstrating the need for an accurate and balanced description of dynamical and non-dynamical correlation effects in highly-quasi-degenerate situations created by the four low-

lying electronic states of  $\text{C}_4\text{O}_4$ . The UB3LYP approach cannot be regarded as a high-level treatment of dynamical correlations, while having potential problems with the non-dynamical ones. CASSCF neglects the non-dynamical correlations altogether. Based on our prior experience with biradical situations, it is expected that CR-CC(2,3),A and, especially, CR-CC(2,3),D provide a more robust, balanced, and accurate treatment of the dynamical and non-dynamical correlations characterizing the  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^3\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$  states of  $\text{C}_4\text{O}_4$ . At the same time, as already alluded to above, we can expect that the EOMCCSD treatment of the open-shell singlet  $9\pi(^1\text{B}_{2u})$  state is reasonably accurate too. As shown in Table 4.11, the differences between the CR-CC(2,3),A and CR-CC(2,3),D geometries, on the one hand, and the UB3LYP and CASSCF geometries, on the other hand, which are reaching a  $0.03 - 0.05 \text{ \AA}$  level for the  $R_{\text{C-O}}$  distance in the  $9\pi(^3\text{B}_{2u})$  and  $10\pi(^1\text{A}_{1g})$  states, are substantial. They are larger than the effect of the basis set on the CCSD, EOMCCSD, and CR-CC(2,3) optimizations and much larger than the differences among the various CC results shown in Table 4.11. The CR-CC(2,3),A and CR-CC(2,3),D geometries are virtually identical (to within  $0.001 \text{ \AA}$ ), independent of the basis set, and, interestingly, not much different than those obtained with the lower-level CCSD approximation (differences on the order of  $0.01 \text{ \AA}$  or smaller if we use the same basis set). This supports one of our conclusions from a different study [180], discussed in Section 3.2, that low-level CC theories are capable of providing reasonable equilibrium geometries compared to their more expensive counterparts with triples, such as CR-CC(2,3). Although for a more reliable description one should use higher-level CC approaches, especially in cases like  $\text{C}_4\text{O}_4$ , where these small differences in geometries can have a significant impact on the final results, it is encouraging to see that

the CCSD geometries are not much different than their CR-CC(2,3) counterparts.

Based on the results in Tables 4.9 and 4.11, we have established that in order to obtain a reliable description of the relative energetics of the  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$  states of  $C_4O_4$ , we need to use basis sets of the aug-cc-pVDZ or better quality and that one has to rely on methods of the CR-CC(2,3) type, i.e., CC methods with a robust treatment of triple excitations, at least for the  $8\pi(^1A_{1g})$ ,  $9\pi(^3B_{2u})$ , and  $10\pi(^1A_{1g})$  states, or methods such as CCSDt and MRCI(Q) with large active orbital spaces. The only state that can be treated at a lower EOMCCSD level is the  $9\pi(^1B_{2u})$  state, regarded as an excitation from the  $8\pi(^1A_{1g})$  state, although even in this case it will at some point become important to examine the role of triples. We have established that it may be reasonable to use the CCSD optimized geometries as long as the basis set used in the calculations is of the aug-cc-pVDZ (or higher) quality, but it is generally ill-advised to perform single-point CC calculations at the geometries obtained with the low-level non-CC methods, such as UB3LYP or CASSCF, without further analysis at the CC levels, since the CC and lower-level non-CC geometry optimizations produce rather different bond lengths that may effect the resulting energetics. It is generally safest to use the geometries optimized at the same CC theory level as that employed in determining the correct relative energetics, since the variation among the results of the single-point calculations is too large if different geometries are employed, especially when one takes into account the small energy spacings among the low-lying states of  $C_4O_4$ . Thus, we are now left with a task of examining the final set of questions, namely, what are the adiabatic energy spacings between the  $9\pi(^3B_{2u})$ ,  $8\pi(^1A_{1g})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$  states of  $C_4O_4$  if we use the high-level CR-CC(2,3) methods with the basis sets of

the aug-cc-pVDZ or better quality and compute the adiabatic energy spacings using the corresponding CR-CC(2,3) geometries and how do our best CR-CC(2,3), $X$ /aug-cc-pVTZ//CR-CC(2,3), $X$ /aug-cc-pVDZ ( $X = A, D$ ) results for the adiabatic gaps involving the  $9\pi(^3B_{2u})$ ,  $8\pi(^1A_{1g})$ ,  $9\pi(^1B_{2u})$ , and  $10\pi(^1A_{1g})$  states compare to the high-level MRCI(Q)(14,13) energetics using CASSCF(16,16)/aug-cc-pVDZ orbitals and reliable CC/aug-cc-pVDZ (rather than unreliable CASSCF(16,16)/6-31G(d)) geometries. Last, but not least, we would like to know how our best CR-CC(2,3), $X$ /aug-cc-pVTZ//CR-CC(2,3), $X$ /aug-cc-pVDZ ( $X = A, D$ ) results for the adiabatic  $8\pi(^1A_{1g}) - 9\pi(^3B_{2u})$  gap compare to the experimentally derived results reported in Refs. [329] and [330].

All of this is examined in Table 4.12. It is clear from this table that our best CR-CC(2,3), $X$ /aug-cc-pVTZ//CR-CC(2,3), $X$ /aug-cc-pVDZ ( $X = A, D$ ) calculations, augmented with the EOMCCSD/aug-cc-pVTZ//EOMCCSD/aug-cc-pVDZ calculations for the open-shell singlet  $9\pi(^1B_{2u})$  state support the state ordering extrapolated from the preliminary results in Table 4.9, namely,  $9\pi(^3B_{2u}) < 8\pi(^1A_{1g}) < 9\pi(^1B_{2u}) < 10\pi(^1A_{1g})$ . In fact, the same ordering is obtained if we use the aug-cc-pVDZ basis set in the energy calculations and geometry optimizations. It is only when we switch to the smaller 6-31G(d) basis set or calculate the  $9\pi(^3B_{2u})$  and  $8\pi(^1A_{1g})$  states using CCSD rather than CR-CC(2,3), when the  $8\pi(^1A_{1g})$  state becomes the ground state. The MRCI(Q)(14,13)/aug-cc-pVDZ results using CASSCF(16,16) references at the CCSD/aug-cc-pVDZ or CR-CC(2,3)/aug-cc-pVDZ geometries are remarkably consistent with our best CR-CC(2,3), $X$ /aug-cc-pVTZ//CR-CC(2,3), $X$ /aug-cc-pVDZ ( $X = A, D$ ) energetics, altering the CR-CC(2,3), $X$ /aug-cc-pVDZ or CR-CC(2,3), $X$ /aug-cc-pVTZ energy spacings at the CR-CC(2,3), $X$ /aug-cc-pVDZ geometries by less than 1

kcal/mol when the  $8\pi(^1A_{1g}) - 9\pi(^3B_{2u})$  gap is examined. The agreement between the MRCI(Q)(14,13)/aug-cc-pVDZ and the corresponding EOMCCSD/aug-cc-pVDZ or EOMCCSD/aug-cc-pVTZ results at the EOMCCSD/aug-cc-pVDZ geometry for the  $9\pi(^1B_{2u})$  state and the agreement between the MRCI(Q)(14,13)/aug-cc-pVDZ and the CR-CC(2,3),X/aug-cc-pVDZ or CR-CC(2,3),X/aug-cc-pVTZ data at the CR-CC(2,3),X/aug-cc-pVDZ geometries ( $X = A, D$ ) for the  $10\pi(^1A_{1g})$  state is not as good as in the case of the  $8\pi(^1A_{1g}) - 9\pi(^3B_{2u})$  gap, but we can rather safely conclude that the  $9\pi(^1B_{2u})$  state is  $\sim 20 - 30$  kJ/mol above the  $8\pi(^1A_{1g})$  state and that the  $10\pi(^1A_{1g})$  state is highest among the four states of  $C_4O_4$  considered here, at about  $30 - 40$  kJ/mol above the  $8\pi(^1A_{1g})$  state. It is quite possible that the MRCI(Q) results would come even closer to our best CC results for the  $9\pi(^1B_{2u})$  and  $10\pi(^1A_{1g})$  states if we could use larger active spaces than (14,13) in designing the MRCI wave functions. It is clear from Table 4.12 that none of the above results are affected by ZPEs, which change the electronic energy spacings by tiny fractions of kJ/mol.

The results in Table 4.12 clearly indicate that the  $9\pi(^3B_{2u})$  state is located  $7 - 11$  kJ/mol below the  $8\pi(^1A_{1g})$  state, in agreement with the experimentally derived values of about  $6 - 7$  kJ/mol reported in Refs. [329] and [330]. It is true that the formally somewhat less complete CR-CC(2,3)-type approach, abbreviated as CR-CC(2,3),A, predicts a  $8\pi(^1A_{1g}) - 9\pi(^3B_{2u})$  gap of  $\sim 7.5$  kJ/mol, in better agreement with experiment than the  $\sim 11.0$  kJ/mol obtained with the formally better CR-CC(2,3),D scheme, but we must keep in mind that the difference between both CR-CC(2,3) methods is less than 1 kcal/mol, i.e., it is within the accuracy these methods can offer. Although we do not want to claim that the CR-CC(2,3),D estimate of the singlet-triplet gap in  $C_4O_4$  is necessarily more accurate than the

CR-CC(2,3),A, MRCI(Q)(14,13), or experimental data, it might be worthwhile to clarify the observed remaining small discrepancies between theory and experiment by carrying out additional NIPE experiments performed on the  $\text{C}_4\text{O}_4$  system, using tricks such as measuring spectra with the laser beams being parallel and perpendicularly polarized, so as to provide more information about the origin of the ejected electrons and a more definitive interpretation of the NIPE spectra, since it has already been demonstrated that issues, such as vibrational progressions and Franck-Condon analyses, may complicate the process of extracting the accurate  $8\pi(^1\text{A}_{1g}) - 9\pi(^3\text{B}_{2u})$  gap value [329, 330].

In summary, we carried out a thorough investigation of the relative ordering of the four lowest-energy densely spaced states of the challenging  $\text{C}_4\text{O}_4$  molecule, abbreviated as  $9\pi(^3\text{B}_{2u})$ ,  $8\pi(^1\text{A}_{1g})$ ,  $9\pi(^1\text{B}_{2u})$ , and  $10\pi(^1\text{A}_{1g})$ , examining the role of the basis set, nuclear geometry relaxation, and level of *ab initio* molecular electronic structure theory employed in the calculations, focusing on the higher-level CC methods and their MRCI-type counterpart. We carried out CCSD, EOMCCSD, CR-CC(2,3),A, CR-CC(2,3),D, CCSDt(8,8), and CCSDt(16,16) single-point energy calculations employing the UB3LYP/6-31G(d), CASSCF(2,2)/6-31G(d), and CASSCF(16,16)/6-31G(d) geometries [327] using the 6-31G(d) and aug-cc-pVxZ ( $x = \text{D}, \text{T}$ ) basis sets, showing that the CR-CC(2,3) calculations are capable of accurately reproducing the high-level CCSDt(16,16) results that are expected to be within  $1 - 2$  kJ/mol from their converged full CCSDT counterparts. We also demonstrated that the geometry relaxation effects and basis set play a significant role in determining the state ordering and the corresponding energy spacings. We thus performed geometry optimizations for all four states of  $\text{C}_4\text{O}_4$  examined here at the CCSD/EOMCCSD, CR-CC(2,3),A, and



CR-CC(2,3),D levels employing the 6-31G(d) and aug-cc-pVDZ basis sets. These geometries were subsequently used to perform various single-point energy calculations using the multireference MRCI(Q)(14,13)/aug-cc-pVDZ level of theory and the CCSD/EOMCCSD, CR-CC(2,3),A, and CR-CC(2,3),D methods employing the larger aug-cc-pVTZ basis set. For the CC/aug-cc-pVTZ//CC/aug-cc-pVDZ results, we computed an approximate vibrational ZPE correction at the UB3LYP/aug-cc-pVTZ level of theory in order to make comparisons with the available experimental data [329]. We demonstrated that one cannot rely on small valence basis sets, such as 6-31G(d), and that basis sets of the aug-cc-pVDZ or better, e.g., aug-cc-pVTZ, quality must be employed to obtain a reasonable description. We also showed that it is not safe to use low-level non-CC geometries, such as those resulting from DFT and CASSCF calculations, in determining the adiabatic energy gaps between the calculated states of  $\text{C}_4\text{O}_4$ . Our best CR-CC(2,3),X/aug-cc-pVTZ//CR-CC(2,3),X/aug-cc-pVDZ ( $X = \text{A, D}$ ) results, supported by the CCSDt/aug-cc-pVDZ and state-averaged MRCI(Q)/aug-cc-pVDZ//CC/aug-cc-pVDZ calculations with large active spaces and the EOMCCSD/aug-cc-pVTZ//EOMCCSD/aug-cc-pVDZ data for the open-shell singlet  $9\pi(^1\text{B}_{2u})$  state, show the ground state of  $\text{C}_4\text{O}_4$  to be the  $9\pi(^3\text{B}_{2u})$  triplet, in agreement with some of the previous theoretical results [325–328, 331] and experiment [329, 330], with the  $8\pi(^1\text{A}_{1g}) - 9\pi(^3\text{B}_{2u})$  singlet-triplet gap being about 7 – 11 kJ/mol. The latter result is in good agreement with the experimental gap estimate of  $6.27 \pm 0.5$  kJ/mol [329], especially if we take into account the challenge of having to deal with a very small energy gap and difficult to balance strong dynamical and nondynamical correlation effects. The state ordering of the four lowest-energy states of  $\text{C}_4\text{O}_4$  is, according to our best CR-CC(2,3), CCSDt, and MRCI(Q) calculations,

$9\pi(^3B_{2u}) < 8\pi(^1A_{1g}) < 9\pi(^1B_{2u}) < 10\pi(^1A_{1g})$ , though further theoretical work may need to determine the more precise location of the  $9\pi(^1B_{2u})$  state. The EOMCCSDt, DEA-EOMCC( $4p-2h$ ), and DIP-EOMCC( $4h-2p$ ) approaches developed by our group or the newer and very promising CC( $P;Q$ ) methodology [68, 150, 151] combining the active-space and CR-CC/CR-EOMCC approaches, developed by our group as well, might be good candidates for this purpose and will be part of future investigations.

# Chapter 5

## Algorithmic advances

In this chapter, the development and application of different algorithms for general use (numerical derivatives) and specific use with the CC/EOMCC routines are discussed. We also discuss the unrestricted implementation of the CR-CC(2,3) methodology, along with the relevant benchmarks.

### 5.1 Parallel numerical derivatives

As shown in the previous chapters, modern advances in *ab initio* electronic structure methods, particularly those based on the CC and EOMCC theories, allow for an accurate treatment and characterization of complex chemical situations, including highly-correlated electronic states, photochemistry, and catalysis. However, one often only uses the high-level CC/EOMCC approaches to verify or benchmark the energetics using the geometries and other molecular properties determined using lower-level approaches, such as DFT, MP2, or CASSCF, as done in some earlier sections. This is not always sufficient, especially when faced with more challenging molecular problems, such as the four lowest-energy states of C<sub>4</sub>O<sub>4</sub> examined in Section 4.4, where determining the energetic ordering of the states required optimizing the geometries using higher-level CC approaches, such as CR-CC(2,3). While analytic gradients and higher-order derivatives would be desirable for the various CC

methods [80, 366–373], such as CR-CC(2,3), and other high-level EOMCC approximations including those in the electron excitation, IP-EOMCC, and EA-EOMCC hierarchies. Their development is often accompanied by its own challenges, especially if one would like to use them in applications to larger chemical species. For example, to determine the analytic gradient of CCSD(T), one has to first compute the  $T_1$  and  $T_2$  clusters at the current unperturbed geometry and, once those have been determined, calculate the energy gradient of CCSD(T), requiring the wall time of approximately two single-point CCSD(T) calculations [369, 370]. These two sequential steps cannot be decoupled and as such are a limiting step in the use of CC/EOMCC analytic gradients. This becomes even more complex in the case of higher-order CC/EOMCC derivatives (cf., e.g., Ref. [374] and references therein for more details). Parallelization of CC/EOMCC analytic derivative codes would help, but this is a lot more involved than in the case of single-point calculations [375].

One way to circumvent some of these issues is through the use of numerical differentiation at the CC/EOMCC levels of theory. Original implementations of numerical gradients for use with high-level CC methods was plagued by having to run the calculations in serial, severely limiting their use and application [376, 377]. With the advent of modern massively-parallel computer architectures this is less of a problem, particularly, since the algorithms to compute numerical derivatives are trivially parallelizable. Yet, in spite of these advances, many of the available quantum chemistry packages still only perform numerical differentiations sequentially making their use costly. For example, to compute a single numerical gradient for a molecular system containing  $M$  atoms one needs  $6M + 1$  single-point energy evaluations

when central finite difference, namely,

$$\frac{\partial f(x_1, \dots, x_i, \dots, x_n)}{\partial x_i} = \frac{f(x_1, \dots, x_i + \frac{1}{2}h_{x_i}, \dots, x_n) - f(x_1, \dots, x_i - \frac{1}{2}h_{x_i}, \dots, x_n)}{h_{x_i}}, \quad (5.1)$$

where  $h_{x_i}$  is the spacing around the variable  $x_i$  and  $f(x_1, \dots, x_n)$  is the function to be differentiated (energy in our case) is employed. For numerical second and higher-order derivatives the situation is even worse, since in the former case  $18M^2 + 1$  single-point calculations are needed (assuming the symmetry of the Hessian has been exploited) as in the expressions below

$$\begin{aligned} \frac{\partial^2 f(x_1, \dots, x_i, \dots, x_j, \dots, x_n)}{\partial x_j \partial x_i} &= \frac{f(x_1, \dots, x_i + \frac{1}{2}h_{x_i}, \dots, x_j + \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} - \\ &\quad \frac{f(x_1, \dots, x_i + \frac{1}{2}h_{x_i}, \dots, x_j - \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} - \\ &\quad \frac{f(x_1, \dots, x_i - \frac{1}{2}h_{x_i}, \dots, x_j + \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} + \\ &\quad \frac{f(x_1, \dots, x_i - \frac{1}{2}h_{x_i}, \dots, x_j - \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}}, \quad (i \neq j), \end{aligned} \quad (5.2)$$

and

$$\begin{aligned} \frac{\partial^2 f(x_1, \dots, x_i, \dots, x_n)}{\partial x_i^2} &= \frac{f(x_1, \dots, x_i + h_{x_i}, \dots, x_n)}{h_{x_i}^2} - \\ &\quad \frac{2f(x_1, \dots, x_i, \dots, x_n)}{h_{x_i}^2} + \\ &\quad \frac{f(x_1, \dots, x_i - h_{x_i}, \dots, x_n)}{h_{x_i}^2}, \quad (i = j). \end{aligned} \quad (5.3)$$

Elementary observations that each of the single-point energy calculations are independent

of one another and that modern computer architectures allow for large-scale parallelization across many nodes with multiple cores implies that even for complex chemical systems numerical gradients and higher-order derivatives can be computed in the wall time of one single-point energy calculation. There clearly is a need for generally applicable numerical derivative routines, which can be used to determine nuclear geometries and other molecular properties at any level of theory, especially for high-level approaches for which even gradients, let alone second- and higher-order derivatives, are very hard to develop and implement. One such program that can do just this is the CIOpt suite [235,378], originally developed by Professors Benjamin G. Levine and Todd J. Martinez to search for minimum-energy crossings without having to compute the derivative couplings, which has been expanded in recent years to evaluate other molecular properties numerically and in parallel, benefitting from the multinode, multicore architectures, where one can run each energy point needed to determine the derivative on a different node, using some or all cores in the node to run this energy calculation in parallel. Our goal has been to interface the CIOpt code to various CC/EOMCC methods, especially those developed by our group, and to extend the CIOpt parallel gradients to second energy derivatives, so that we can determine harmonic vibrational frequencies through parallel numerical differentiation.

### 5.1.1 Gradients for geometry optimizations

Using the embarrassingly parallel nature of numerical derivative algorithms as implemented in CIOpt, we optimized the geometries of a variety of molecules that had originally been optimized using serial numerical gradients available in GAMESS with the CCSD, CR-CC(2,3),

EOMCCSD, EA-EOMCC, and IP-EOMCC approaches as discussed in previous sections of this thesis, using several different architectures to test the scalability of these routines. The results and timings of these calculations along with the larger scale optimizations discussed in Section 5.1.3 below are shown in Table 5.1 and Table 5.3 in Section 5.1.3. As can be seen from these tables, the numerical gradient routines as implemented in CIOpt show almost perfect scalability, at least for problems up to the same size as those we were able to test. For example, when azulene, which has 18 degrees of freedom when the  $C_{2v}$  symmetry is employed, was optimized using 37 nodes (the total number of single points needed to compute one gradient is  $2 \times 18 + 1 = 37$ ) a speedup of  $\sim 35$  was observed. While this is not quite a perfect speedup, where the factor would be exactly 37, this is very encouraging showing very little loss in efficiency as larger numbers of cores are employed to compute a numerical gradient. In all of the optimizations reported in previous sections of this thesis, we used the BFGS algorithm [379] as included in the GAMESS and CIOpt programs, with the latter implementation discussed below.

Table 5.1: Wall time (hours) comparison of parallel and serial numerical geometry optimizations of several molecules.

| Molecular                     |                 |                 |                   |            |                    | Cores/ Time/Gradient |                     |                    |         |         |
|-------------------------------|-----------------|-----------------|-------------------|------------|--------------------|----------------------|---------------------|--------------------|---------|---------|
| System                        | Symmetry        |                 | Method            | Basis      | Degrees of Freedom | Nodes                | Node                | Parallel           | Serial  | Speedup |
| C <sub>2</sub> N <sup>a</sup> | C <sub>2v</sub> |                 | EA-EOMCCSDt{4}    | cc-pVQZ    | 2                  | 1                    | 5                   | 1.49               | 5.68    | 3.81    |
| C <sub>2</sub> N <sup>a</sup> | C <sub>2v</sub> |                 | EA-EOMCCSD(3p-2h) | cc-pVTZ    | 2                  | 1                    | 5                   | 8.47               | 42.37   | 5.00    |
| Azulene <sup>a</sup>          | C <sub>2v</sub> |                 | CR-CC(2,3),A      | cc-pVDZ    | 18                 | 37                   | 1                   | 55.27              | 1916.53 | 34.7    |
| Au <sub>8</sub> <sup>b</sup>  | S <sub>1</sub>  | D <sub>2h</sub> | CCSD(T)           | SBKJC(1f)  | 4                  | 9                    | 8                   | 9.57               | -       | -       |
| Au <sub>8</sub> <sup>b</sup>  | S <sub>1</sub>  | D <sub>4h</sub> | CCSD(T)           | SBKJC(1f)  | 2                  | 5                    | 8                   | 9.57               | -       | -       |
| Au <sub>8</sub> <sup>b</sup>  | S <sub>3</sub>  | T <sub>d</sub>  | CCSD(T)           | SBKJC(1f)  | 4                  | 9                    | 8                   | 10.39              | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>4</sub>  | C <sub>s</sub>  | CCSD(T)           | SBKJC(1f)  | 7                  | 2                    | 24, 16 <sup>c</sup> | 12.84 <sup>g</sup> | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>6</sub>  | D <sub>2d</sub> | CCSD(T)           | SBKJC(1f)  | 4                  | 2                    | 24, 16 <sup>c</sup> | 13.20              | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>1</sub>  | D <sub>4h</sub> | CCSD(T)           | cc-pVDZ-PP | 2                  | 2                    | 27,18 <sup>d</sup>  | 17.55              | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>3</sub>  | T <sub>d</sub>  | CCSD(T)           | cc-pVDZ-PP | 4                  | 2                    | 30,24 <sup>e</sup>  | 26.82              | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>4</sub>  | C <sub>s</sub>  | CCSD(T)           | cc-pVDZ-PP | 7                  | 2                    | 32,28 <sup>f</sup>  | 41.90              | -       | -       |
| Au <sub>8</sub> <sup>a</sup>  | S <sub>6</sub>  | D <sub>2d</sub> | CCSD(T)           | cc-pVDZ-PP | 4                  | 2                    | 30,24 <sup>e</sup>  | 28.90              | -       | -       |

<sup>a</sup>This calculation was run on nodes consisting of four eight-core Intel Xeon X7560s at 2.26 GHz with 256 GB of RAM.

<sup>b</sup>This calculation was run on nodes consisting of two four-core Intel Xeon E5620s (Westmere family) at 2.4 GHz with 24 GB of RAM at Michigan State University's High Performance Computing Center. These nodes were about 30-50% faster than the other nodes used to run computations.

<sup>c</sup>Each single-point CCSD(T) calculation was run using 8 cores.

<sup>d</sup>Each single-point CCSD(T) calculation was run using 9 cores.

<sup>e</sup>Each single-point CCSD(T) calculation was run using 6 cores.

<sup>f</sup>Each single-point CCSD(T) calculation was run using 4 cores.

<sup>g</sup>Each gradient was computed in three groups of five single-point CCSD(T) calculations. The time shown is the average time is took to run each of the three groups to compute a gradient, for all optimization cycles.



As mentioned above, numerical gradients, or first derivatives, can be used for determining molecular properties, such as equilibrium geometries. Pulay was the first to demonstrate their efficient use in applications to chemical systems [380] and since then much work has been done to progress the field. There are several approaches which seek to avoid the explicit computation of the Hessian [381] (and in some cases even the gradient) in the optimization of a function (energy in our considerations in this thesis) including the conjugate gradient approach [382], Nelder-Mead or downhill simplex method [383], or Quasi-Newton approximations. In the latter approach, the explicit computation of the Hessian is replaced by recurring at each optimization cycle an approximate Hessian, allowing the curvature of the problem to be exploited implicitly.

Among the various Quasi-Newton algorithms which have been developed one of the more popular is the BFGS method [384–388], named for its discoverers Profs. C.G. Broyden, R. Fletcher, D. Goldfarb, and D.F Shanno. The BFGS algorithm employing numerical gradients is outlined as follows:

1. Given the starting point, or geometry in our case,  $x_0$ , convergence tolerance  $\epsilon > 0$ , and an approximate inverse Hessian  $H_0$  compute the the gradient  $\nabla f_k$  of the function being minimized  $f$ , where  $k$  is the number of the current optimization cycle.
2. If  $\|\nabla f_k\| > \epsilon$  compute the search direction  $d_k = -H_k \nabla f_k$  and the new set of parameters  $x_{k+1} = x_k + \alpha_k d_k$  where  $\alpha_k$  is a suitably chosen parameter with either a fixed or variable value, depending on the implementation and can be determined from, for example, a line-search or trust-region calculation as the minimizer of  $f(x_k + \alpha_k d_k)$ .
3. Update the approximate inverse Hessian,  $H_k$  using the following, or one of its other

equivalent variations

$$H_{k+1} = (\mathbf{1} - \rho_k s_k y_k^T) H_k (\mathbf{1} - \rho_k y_k s_k^T) + \rho s_k s_k^T \quad (5.4)$$

where  $s_k = x_{k+1} - x_k$ ,  $y_k = \nabla f_{k+1} - \nabla f_k$  and  $\rho = (y_k^T s_k)^{-1}$ .

4. Iterate until  $\|\nabla f_k\| < \epsilon$ .

One of the reasons for its popularity is that it preserves the symmetric and (semi) positive definite properties of  $H_k$  (e.g.,  $x^T H_k x > 0, \forall x \in \mathbb{R}^n$ , since  $H_k$  is symmetric), such that its eigenvalues remain positive and real throughout the optimization process. The symmetry and positive definiteness are preserved through the use of a rank-two update, Eq. (5.4).

A simpler rank-one matrix update to  $H_k$ , which maintains its symmetry though not necessarily its positive definiteness, was proposed in Ref. [389–391] called the symmetric rank-one (SR1) approach. It has been shown that in many cases this version of the SR1 algorithm outperforms the BFGS method [389–392] and that, unlike with BFGS, the sequence of matrices generated by the SR1 process  $H_k$  converge to the true Hessian  $H_*$  under certain conditions [391]. In spite of the many examples showing that the SR1 algorithm rarely violates the positive definiteness of the Hessian with its update, it has failed to garner as much use or attention as the BFGS approach. In geometry optimizations of chemical systems there has appeared one study, in which the SR1 approach was employed [393] using the Hartree-Fock level of theory. Particularly for the molecules with a larger number of degrees of freedom, it was shown that the SR1 method can drastically reduce the number of optimization cycles required to locate a stationary point corresponding to a minimum on

the potential energy surface of the molecule. There has been no followup by the author of Ref. [393] and no one else has pursued this topic in terms of geometry optimizations for chemical species and systems.

While the originally proposed SR1 algorithm [389–391] that interests us in this thesis work does not preserve the positive definite nature of  $H_k$  it has been shown that deviations from positive definiteness in the approximate update to  $H_k$  are related to the condition number of the matrix and positive definiteness can therefore be restored using an optimal scaling factor [394, 395]. Using this idea, the authors of Ref. [396] proposed what they called a scaled “memoryless” SR1 method, though it is actually a limited-memory (LM) algorithm as shown below, in which the positive definite nature of  $H_k$  is preserved throughout the optimization. In order to make their SR1 approach a LM method they used the simple trick of replacing the inverse Hessian  $H_k$  by the identity matrix  $\mathbf{1}$  multiplied by an appropriate scaling factor  $\gamma_k$ , which results in the original update

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{y_k^T (s_k - H_k y_k)} \quad (5.5)$$

simplifying to

$$H_{k+1} = \gamma_k \mathbf{1} + \frac{(s_k - \gamma_k y_k)(s_k - \gamma_k y_k)^T}{y_k^T (s_k - \gamma_k y_k)} \quad (5.6)$$

where

$$\gamma_k = \frac{s_k^T s_k}{s_k^T y_k} - \left[ \left( \frac{s_k^T s_k}{s_k^T y_k} \right)^2 - \frac{s_k^T s_k}{y_k^T y_k} \right]^{1/2} \quad (5.7)$$

and instead of having to store the inverse Hessian matrix of size  $n \frac{(n+1)}{2}$  one only need store

Table 5.2: Comparison of the LM-SR1 and BFGS quasi-Newton algorithms for CR-CC(2,3),D/TZVP geometry optimizations.<sup>a</sup>

| Molecule (symmetry)           | DoF <sup>b</sup> | Optimization Cycles |      |
|-------------------------------|------------------|---------------------|------|
|                               |                  | LM-SR1              | BFGS |
| Benzene ( $D_{6h}$ )          | 2                | 9                   | 7    |
| Cyclopentadiene ( $C_{2v}$ )  | 7                | 10                  | 12   |
| Acetone ( $C_{2v}$ )          | 9                | 7                   | 14   |
| Formamide ( $C_s$ )           | 12               | 3                   | 15   |
| E-Butadiene ( $C_{2h}$ )      | 10               | 12                  | 24   |
| all-E-Hexatriene ( $C_{2h}$ ) | 14               | 20                  | 50 < |
| Pyridine ( $C_{2v}$ )         | 11               | 8                   | 41   |
| Thymine ( $C_s$ )             | 29               | 6                   | 28   |

<sup>a</sup> As many cores as single-point energies needed to compute the numerical gradient were used (2 DoF + 1).

<sup>b</sup> Number of degrees of freedom (DoF) optimized

the several vectors each of size  $n$ . The step direction  $d_k = -H_k \nabla f_k = -H_k g_k$  then becomes

$$\begin{aligned}
 d_k = & -\gamma_{k-1} g_k + \gamma_{k-1} \left( \frac{s_{k-1}^T g_k - \gamma_{k-1} y_{k-1}^T g_k}{y_{k-1}^T s_{k-1} - \gamma_{k-1} y_{k-1}^T y_{k-1}} \right) y_{k-1} \\
 & - \left( \frac{s_{k-1}^T g_k - \gamma_{k-1} y_{k-1}^T g_k}{y_{k-1}^T s_{k-1} - \gamma_{k-1} y_{k-1}^T y_{k-1}} \right) s_{k-1}.
 \end{aligned} \tag{5.8}$$

Using the same general outline as for the BFGS algorithm above with the  $H_k$  and  $d_k$  updates being replaced by those in Eqs. (5.6) and (5.8) the authors of Ref. [396] were able to show that their LM-SR1 algorithm was able to perform at least as well, if not better, than the standard BFGS and LM-BFGS [397] methods.

Encouraged by the previous study [393] and these improvements [396] to the SR1 algorithm, the LM-SR1 and a version of the LM-BFGS approaches were implemented in the CIOpt program. The LM-SR1 and BFGS algorithm from Ref. [379] were used in conjunc-

tion with the parallel numerical gradient routine also available in CIOpt and interfaced with GAMESS to optimize the geometry of 8 molecules of varying size, namely, benzene, cyclopentadiene, acetone, formamide, E-butadiene, all-E-hexatriene, pyridine, and the nucleobase thymine at the CR-CC(2,3),D level of theory using the TZVP basis [398]. The results of these equilibrium geometry optimizations are shown in Table 5.2. The gradients were converged to  $10^{-4}$  and the difference in energy for the SR1 and BFGS optimized structures were less than  $10^{-5}$  in all cases. For benzene where we employed the  $D_{6h}$  spatial symmetry and only two degrees of freedom (DoF) were optimized the BFGS approach reaches the minimum in fewer optimization cycles than the LM-SR1 method. When the number of DoF becomes larger, as in the case of thymine, we see a more pronounced difference between the two algorithms. Probably the more impressive case is that of all-E-hexatriene where the LM-SR1 approach takes 20 optimization cycles to reach a minimum, while the BFGS method took more than 50 for the same system. The nucleobase thymine, which has a large number of DoF (29), particularly when optimizing the geometry with such high-level approaches as CR-CC(2,3), where our LM-SR1 algorithm performs quite well, reaching the minimum energy nuclear configuration in just six cycles whereas the equivalent BFGS optimization took 28 cycles. Clearly more studies need to be carried out, but these initial results are quite encouraging. The LM-SR1 algorithm has also been used on a couple of test cases, namely ammonia and ethylene, using the  $\delta$ -CR-EOMCC(2,3),D and CASSCF approaches, respectively, to evaluate its performance in locating minimum energy crossings between different electronic states of a given molecule with encouraging results, which will be followed up in future studies.

### 5.1.2 Hessians for harmonic vibrational analyses

As mentioned above, higher-order numerical derivatives beyond gradients can become very expensive, especially when run in serial. On the other hand, as seen by the lack of analytic Hessians and higher-order analytic derivatives for high-level *ab initio* CC/EOMCC approaches, efficient parallel numerical second and higher derivatives for use with any level of theory has the potential to open new avenues for predicting and interpreting various experimental results. As part of the work done for this dissertation, an efficient parallel numerical second derivative routine for computing harmonic vibrational frequencies was developed and implemented in the CIOpt package. In implementing the numerical second derivative algorithm we made use of the fact that the Hessian is a symmetric matrix whose entries are

$$H_{ij} = \left( \frac{\partial E}{\partial x_i \partial x_j} \right) = \left( \frac{\partial E}{\partial x_j \partial x_i} \right), \quad (5.9)$$

where  $E$  is the energy and  $x_i$  and  $x_j$  represent the degrees of freedom. By using this symmetry we need only compute  $M \frac{(M+1)}{2}$  entries rather than the full  $M^2$  Hessian  $H$  (note: we are using  $H$  to represent the Hessian whereas in the previous section we used it to denote the inverse Hessian; in Section 5.1 this is done in keeping with the standard notation used in numerical analysis where the Hessian itself is represented by  $B$ . In this section we do not use  $B$  for the Hessian, but rather  $H$  so as to avoid confusion with Wilson's  $B$ -matrix discussed below). To parallelize this portion of the algorithm it was decided that flattening the two-dimensional Hessian to one dimension would be the most efficient as it would allow for use of the MPI routines already available in CIOpt with minimal modification. To flatten the entire  $H$  matrix, which is an  $M \times M$  matrix, one uses the trivial formula where the two-

index element  $H_{ij}$  is mapped to one index with  $k = j - 1 + (i - 1)M$ , where  $i$  denotes the row and  $j$  the column of  $H$ , where  $H_{ij}$  resides. However, where we are only interested in those elements along and above/below the diagonal the mapping formula has to be slightly modified for  $i > 1$  such that

$$k = j + (i - 1)M - i \frac{(i + 1)}{2}, \quad (5.10)$$

where the last term on the right-hand side accounts for the fact that we are subtracting the lower-diagonal indices of an  $i \times i$  sub-matrix. Using this same idea, the higher-order derivatives, such as the third and fourth derivatives, can be more easily parallelized as well. For example, the three indices of the third derivative  $i, j$ , and  $k$  can be flattened into a unique index via

$$\ell = k + (j - 1)M - j \frac{(j - 1)}{2} + \frac{1}{2} \sum_{\alpha=2}^i [M(M - 2\alpha + 3) + (\alpha - 1)(\alpha - 2)], \quad (5.11)$$

and for fourth derivatives we have

$$m = \ell + (k - 1)M + (j - 1)M^2 + (i - 1)M^3 - \sum_{\beta=M-i+2}^M \sum_{\alpha=1}^i (\alpha - 1)\beta \frac{(\beta + 1)}{2}. \quad (5.12)$$

These will be useful for future work as the third and fourth derivatives are needed to compute various spectroscopic properties beyond the Harmonic approximation.

For each of the diagonal elements of  $H$  we need to compute two unique single-point

energies

$$\begin{aligned} \frac{\partial^2 E(x_1, \dots, x_i, \dots, x_n)}{\partial x_i^2} &= \frac{E(x_1, \dots, x_i + h_{x_i}, \dots, x_n)}{h_{x_i}^2} - \\ &\quad \frac{2E(x_1, \dots, x_i, \dots, x_n)}{h_{x_i}^2} + \\ &\quad \frac{E(x_1, \dots, x_i - h_{x_i}, \dots, x_n)}{h_{x_i}^2}, \quad (i = j) \end{aligned} \quad (5.13)$$

as well as the energy at the unperturbed/equilibrium geometry. For each of the off-diagonal elements we need four unique single-point energies

$$\begin{aligned} \frac{\partial^2 E(x_1, \dots, x_i, \dots, x_j, \dots, x_n)}{\partial x_j \partial x_i} &= \frac{E(x_1, \dots, x_i + \frac{1}{2}h_{x_i}, \dots, x_j + \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} - \\ &\quad \frac{E(x_1, \dots, x_i + \frac{1}{2}h_{x_i}, \dots, x_j - \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} - \\ &\quad \frac{E(x_1, \dots, x_i - \frac{1}{2}h_{x_i}, \dots, x_j + \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}} + \\ &\quad \frac{E(x_1, \dots, x_i - \frac{1}{2}h_{x_i}, \dots, x_j - \frac{1}{2}h_{x_j}, \dots, x_n)}{h_{x_j}h_{x_i}}, \quad (i \neq j). \end{aligned} \quad (5.14)$$

Thus there are a total of  $2M + 4M\frac{(M-1)}{2} + 1 = 2M^2 + 1$  single-points needed to evaluate the Hessian  $H$  numerically. When discussing the numerical derivatives, particularly second and higher, it must be remembered that these are meant to be run in parallel and not sequentially, since, for example, if we consider a triatomic system, which has 9 degrees of freedom in the three dimensional Cartesian space, then a total of 162 single-point energies need to be computed. It is well known that for nonlinear (linear) molecules there are only



$M - 6$  ( $M - 5$ ) degrees of freedom that need to be considered for harmonic vibrational frequency analysis, we can see that this translates into, for the case of a triatomic molecule, only 18 (32) single-point energy computations. There are several approaches for reducing the DoF of a molecular system to the minimal number [399–406], all of which make use of a variant of Wilson’s  $B$ -matrix [407]. Implementing a routine that generates a form of Wilson’s  $B$ -matrix for efficient parallel numerical second derivatives has been started and will be finished as part of future work.

To evaluate the harmonic vibrational frequencies once the numerical Hessian has been computed, following Ref. [407], the Hessian is mass weighted by multiplying each element  $H_{ij}$  by  $1/(m_i m_j)^{1/2}$  where  $m_i$  is the mass of atom  $i$ . The mass-weighted Hessian is then diagonalized producing the set of normal modes  $Q_\alpha$  and mass-weighted force-constants  $\kappa_\alpha$  ( $\alpha = 1, \dots, M$ ), where the latter are used to compute the harmonic vibrational frequencies using

$$\nu_\alpha = \frac{1}{2\pi} \sqrt{\kappa_\alpha}. \quad (5.15)$$

An algorithm to determine the Abelian symmetry of the vibrational frequencies was implemented, where the irreducible representation to which the normal mode belonged was determined using [408]

$$\mathfrak{P}_q Q_\alpha = \frac{\ell_q}{h} \sum_{\mathfrak{R}} \chi^q(\mathfrak{R}) P_{\mathfrak{R}} Q_\alpha, \quad (5.16)$$

where  $\mathfrak{P}_q$  are the projection operators of the point group representation,  $\ell_q$  is the dimension of the irreducible representation of the group,  $h$  is the dimension of the point group,  $\chi^q(\mathfrak{R})$  are the characters for a given irreducible representation  $q$  of the group,  $\mathfrak{R}$  are the operations of the point group, and  $P_{\mathfrak{R}}$  are the operators/matrix representations corresponding

to  $\mathfrak{R}$ . If the above sum is non-zero then the  $Q_\alpha$  normal mode belongs to that irreducible representation. This formula as written only applies to Abelian point groups. For non-Abelian groups the situation becomes much more complicated whereas one can compute the harmonic vibrational frequencies using an Abelian symmetry group and then using correlation tables determine the symmetry of the normal mode frequencies in the higher-symmetry non-Abelian group.

The above algorithms have been implemented in the CIOpt program package and have been tested on a small number of test cases showing promising results, though more studies need to be carried out with more examples demonstrating their utility.

### 5.1.3 Application to the low-energy isomers of $\text{Au}_8$

The discovery that small ( $2\text{nm} \leq \text{diameter} \leq 4\text{nm}$ ) gold particles  $\text{Au}_n$  can selectively catalyze chemical reactions, as already discussed in Sections 3.3 and 4.3, such as the epoxidation of propene [409], has inspired a lot of activity among experimentalists and theorists toward understanding the origins of this catalytic behavior. Several factors, including surface roughening, may play an important role in the catalytic activity of  $\text{Au}_n$  clusters, since non-planarity of the clusters localizes the electron density and promotes reactivity [410]. Because of the importance of surface roughening in the catalytic activity of gold clusters, it is essential to determine the number of gold atoms in the  $\text{Au}_n$  particle for which the planar-to-non-planar turnover occurs and the non-planar isomers begin to dominate as the lowest-energy species. In the earlier study by Olsen *et al.* [214], calculations using the CCSD(T) approach suggested that the most stable structures of  $\text{Au}_n$  were planar for  $n = 6$  and non-planar for

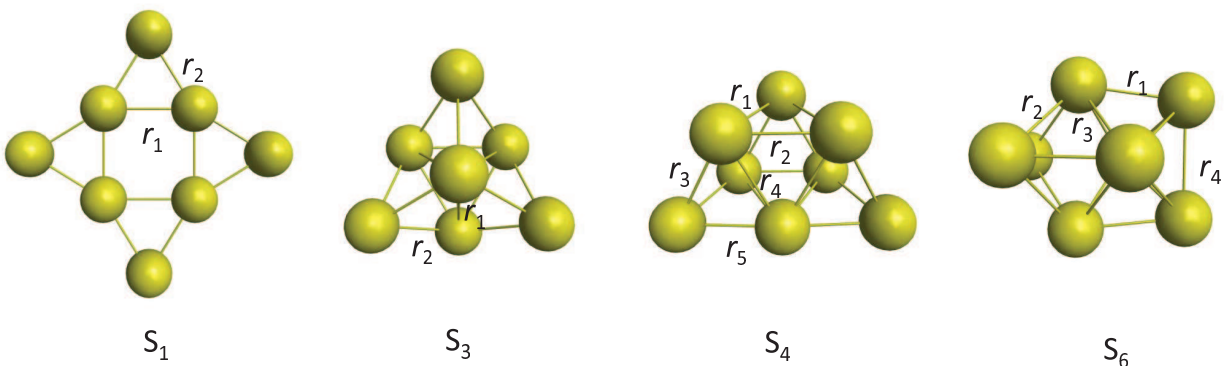


Figure 5.1: The  $S_1$ ,  $S_3$ ,  $S_4$ , and  $S_6$  isomers of  $\text{Au}_8$  examined in this study, along with the selected geometrical parameters included in Table 5.4.

$n = 8$ , in contrast with DFT calculations, which typically predict  $\text{Au}_8$  to favor the planar configuration (cf., e.g., Refs. [204–207], and references therein; see, also, the introduction to Ref. [214] for additional remarks).

When the initial CCSD(T) calculations for  $\text{Au}_8$  were reported [214], issues, such as the use of larger basis sets, geometry relaxation, and the numbers of correlated electrons used in CC calculations could not be addressed due to prohibitive computational costs. Since then several other CC, MP2, and DFT calculations have appeared in the literature [215–217, 411–414], usually in disagreement with the findings in Ref. [214], including the CCSD(T) calculations with larger basis sets and some core correlations that lean toward the conclusion that the lowest-energy  $\text{Au}_8$  isomer should indeed be planar [215–217].

However, the topic still remains open, since all CC calculations for  $\text{Au}_8$  to date rely on low-order methods, such as MP2 or DFT, which do not necessarily provide the correct energetics, to determine nuclear geometries. There also are indications that clusters with 7 or more gold atoms may be non-planar [415]. In view of all this and as part of this thesis work we

performed new CCSD(T) calculations for Au<sub>8</sub> employing larger basis sets than those used in the initial calculations [214] and examining the role of core-valence (CV) correlations, as has been done in Refs. [215–217], and addressing the issue of geometry relaxation at the same time by reoptimizing the low-energy structures of the Au<sub>8</sub> particle using CCSD(T) [183]. Clearly, the issue of geometry relaxation is a challenging one at the CC theory level due to high computational costs, but, as demonstrated in this dissertation, we can take advantage of the efficient parallel numerical gradients for geometry optimizations, which can work with any CC method, including the parallel implementation of CCSD(T) utilized in this study.

As part of the work for the completion of this dissertation, we undertook a systematic investigation of the role of geometry optimization, basis set, and semi-core (i.e., CV) electron correlations in the CCSD(T) calculations of the relative energetics of a few lowest-energy Au<sub>8</sub> isomers in the hope of providing a more definitive answer as to whether or not the most stable structure of Au<sub>8</sub> might be non-planar. We focus on the four lowest-energy isomers of Au<sub>8</sub>, shown in Fig. 5.1, as predicted in Ref. [214] by the small basis set MP2 geometry optimizations and CCSD(T) single-point calculations, reexamined later in Ref. [217]. For each of the four structures, two sets of geometry optimizations, correlating in each case the  $5d^{10}6s^1$  valence electrons, were carried out. In the first set, further elaborated on below, we used the parallel analytic gradients of MP2, available in the GAMESS package [190, 191]. In the second set of geometry optimizations, further discussed below as well, we used the dual-level parallelism applied to the numerical CCSD(T) derivatives, in which the coarse-grain finite-difference model available in the CIOpt program suite [235, 378], utilizing as many nodes as the number of energies needed to determine the gradient, was combined with

the fine-grain single-point CCSD(T) calculations on each node. In running the required single-point CCSD(T) computations, we used the highly scalable parallel CCSD(T) codes, available in GAMESS and developed in Refs. [416,417], which are based on the CCSD(T) algorithm from Ref. [238]. By combining the embarrassingly parallel finite-difference model with the highly scalable CCSD(T) approach, as describe above, we were able to determine a single energy gradient when employing the CCSD(T)/SBKJC(f) approach in about 9 hours when running on nine nodes each with 8 cores (cf. Table 5.3), while at the CCSD(T)/cc-pVDZ-PP level in  $\sim 41$  hrs, when 4 cores per node were used, or  $\sim 18$  hrs, when 9 cores were utilized in each of the required single-point calculations. The difference in timings between the two bases reflects on the difference in their size and quality with the larger cc-pVDZ-PP being the better of the two, as shown in the results below. The numbers of the single-point calculations needed to determine a single energy gradient were 5 for the  $D_{4h}$ -symmetric  $S_1$  structure (9 when run as a  $D_{2h}$  configuration), 9 for the  $T_d$ -symmetric  $S_3$  and  $D_{2d}$ -symmetric  $S_6$  isomers, and 15 for the  $C_s$ -symmetric  $S_4$  configuration. By using the previously converged MP2 geometries, we typically needed 5–10 CCSD(T) optimization cycles to converge the gradient to  $10^{-4}$ .

We started our examination of the MP2 and CCSD(T) energetics using the MP2-optimized geometries. The small basis set MP2 geometry optimizations using the standard SBKJC basis and the associated scalar relativistic ECP, which was optimized for gold [418], augmented by a single set of f functions (exponent = 0.89), referred to as SBKJC(1f), were carried out in Ref. [214]. The resulting CCSD(T)/SBKJC(1f)//MP2/SBKJC(1f) energies have led to the aforementioned conflicting findings about the ordering of the various structures of Aug.

Table 5.3: Details of the wall times (hours) characterizing the CCSD(T)/SBKJC(1f) parallel numerical optimization of the Au<sub>8</sub> S<sub>1</sub> isomer employing the D<sub>2h</sub> symmetry (4 degrees of freedom), performed using nine 8-core nodes<sup>a</sup>.

| Optimization Cycle | Node Number |      |      |      |      |      |       |      |      | Average |
|--------------------|-------------|------|------|------|------|------|-------|------|------|---------|
|                    | 0           | 1    | 2    | 3    | 4    | 5    | 6     | 7    | 8    |         |
| 1                  | 9.67        | 9.60 | 9.50 | 9.52 | 9.80 | 9.82 | 9.82  | 9.52 | 9.48 | 9.64    |
| 2                  | 9.43        | 9.53 | 9.42 | 9.30 | 9.42 | 9.58 | 9.85  | 9.43 | 9.90 | 9.54    |
| 3                  | 9.80        | 9.63 | 9.53 | 9.60 | 9.80 | 9.62 | 10.02 | 9.80 | 9.57 | 9.71    |
| 4                  | 9.87        | 9.77 | 9.67 | 9.72 | 9.70 | 9.65 | 10.10 | 9.67 | 9.58 | 9.75    |
| 5                  | 9.42        | 9.48 | 9.48 | 9.48 | 9.42 | 9.58 | 9.83  | 9.48 | 9.43 | 9.51    |
| 6                  | 9.37        | 9.63 | 9.37 | 9.35 | 9.37 | 9.38 | 9.68  | 9.42 | 9.40 | 9.44    |
| 7                  | 9.35        | 9.40 | 9.37 | 9.37 | 9.37 | 9.40 | 9.70  | 9.40 | 9.35 | 9.41    |
| 8                  | 9.40        | 9.72 | 9.45 | 9.43 | 9.43 | 9.50 | 10.00 | 9.62 | 9.35 | 9.54    |
| Average            | 9.54        | 9.60 | 9.47 | 9.47 | 9.54 | 9.57 | 9.88  | 9.54 | 9.51 | 9.57    |

<sup>a</sup>This calculation was run on nodes consisting of two four-core Intel Xeon E5620s (Westmere family) at 2.4 GHz with 24 GB of RAM at Michigan State University’s High Performance Computing Center.

Thus, in order to examine the role of the basis set used to determine the MP2 geometries, we reoptimized all four Au<sub>8</sub> structures shown in Fig. 5.1 using the MP2 method and the cc-pVDZ-PP basis [229] combined with the scalar relativistic ECP created for use with the correlation-consistent basis sets, which was originally developed in Ref. [230] and later modified to include the effects of the  $h$  functions in the pseudopotential in Ref. [231]. Using these improved MP2/cc-pVDZ-PP geometries, single-point CCSD(T) calculations were carried out exploiting the cc-pV $x$ Z-PP ( $x = D, T$ ) basis sets, combined once again with the ECP of Refs. [230, 231], correlating the same number of electrons corresponding to the  $5d$  and  $6s$  shells of the gold atoms as in the geometry optimizations. To investigate the role of the CV correlation effects, the more extensive CCSD(T)/cc-pV $x$ Z-PP+CV ( $x = D, T$ ) calculations correlating the  $5s^25p^6$  semi-core and  $5d^{10}6s^1$  valence electrons were performed as well. Throughout this section and as done in other portions of this dissertation, we use the notation in which the cc-pV $x$ Z-PP acronym refers to the calculations combining the

cc-pV $x$ Z-PP basis set [229] and the ECP of Refs. [230, 231]. If the  $5s^25p^6$  electrons are correlated in addition to the valence  $5d^{10}6s^1$  electrons, the corresponding computation is abbreviated (following Ref. [217]) as cc-pV $x$ Z-PP+CV.

Having examined the role of the basis set on the MP2 geometries, we went to the next step in which we performed the CCSD(T) geometry optimizations employing the SBKJC(1f) and cc-pVDZ-PP bases, and the accompanying ECPs. In this way, we could examine the significance of higher-order correlation effects on the calculated geometries and the effect of the basis set on the geometry optimizations at the CCSD(T) level of theory. Single-point CCSD(T)/cc-pV $x$ Z-PP and CCSD(T)/cc-pV $x$ Z-PP+CV ( $x = D, T$ ) calculations, using, once again, the relativistic ECPs from Refs. [230, 231], were then carried out to determine the final relative energetics of the four Au<sub>8</sub> structures shown in Fig. 5.1. We have not considered the effects of relativity on the calculated geometries and energies beyond those captured by ECPs, such as spin-orbit coupling, since it is well established that the spin-orbit coupling does not alter the relative stability of Au <sub>$n$</sub>  clusters with  $n \leq 20$ , while having negligible effect on their geometries [419] (see, also, Refs. [215, 411, 412]).

Table 5.4: Bond lengths (in Å) of the S<sub>1</sub>, S<sub>3</sub>, S<sub>4</sub>, and S<sub>6</sub> isomers of Au<sub>8</sub> obtained from geometry optimizations using MP2, CCSD(T), and some representative DFT approaches (see Fig. 5.1 for the meaning of the geometrical parameters).

| Method                        | Basis      | S <sub>1</sub>        |                       | S <sub>3</sub>        |                       | S <sub>4</sub>        |                       |                       |                       |                       | S <sub>6</sub>        |                       |                       |                       |
|-------------------------------|------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                               |            | <i>r</i> <sub>1</sub> | <i>r</i> <sub>2</sub> | <i>r</i> <sub>1</sub> | <i>r</i> <sub>2</sub> | <i>r</i> <sub>1</sub> | <i>r</i> <sub>2</sub> | <i>r</i> <sub>3</sub> | <i>r</i> <sub>4</sub> | <i>r</i> <sub>5</sub> | <i>r</i> <sub>1</sub> | <i>r</i> <sub>2</sub> | <i>r</i> <sub>3</sub> | <i>r</i> <sub>4</sub> |
| MP2 <sup>a</sup>              | SBKJC(1f)  | 2.697                 | 2.592                 | 2.804                 | 2.672                 | 2.704                 | 2.654                 | 2.686                 | 2.758                 | 2.612                 | 2.655                 | 2.703                 | 2.762                 | 2.779                 |
| CCSD(T) <sup>a</sup>          | SBKJC(1f)  | 2.771                 | 2.643                 | 2.865                 | 2.743                 | 2.766                 | 2.711                 | 2.755                 | 2.832                 | 2.674                 | 2.714                 | 2.772                 | 2.819                 | 2.908                 |
| MP2 <sup>a</sup>              | cc-pVDZ-PP | 2.707                 | 2.588                 | 2.885                 | 2.677                 | 2.704                 | 2.672                 | 2.693                 | 2.766                 | 2.612                 | 2.664                 | 2.713                 | 2.778                 | 2.810                 |
| CCSD(T) <sup>a</sup>          | cc-pVDZ-PP | 2.782                 | 2.643                 | 2.896                 | 2.751                 | 2.768                 | 2.723                 | 2.766                 | 2.850                 | 2.676                 | 2.728                 | 2.783                 | 2.829                 | 2.964                 |
| LC- $\omega$ PBE <sup>b</sup> | cc-pVDZ-PP | 2.759                 | 2.626                 |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |
| $\omega$ B97X <sup>b</sup>    | cc-pVDZ-PP | 2.817                 | 2.663                 |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |
| TPSS <sup>b</sup>             | cc-pVDZ-PP | 2.758                 | 2.633                 |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |
| CAM-B3LYP <sup>b</sup>        | cc-pVDZ-PP | 2.806                 | 2.648                 |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |
| B3LYP <sup>b</sup>            | cc-pVDZ-PP | 2.830                 | 2.672                 |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |

<sup>a</sup>This work. Although the MP2/SBKJC(1f) geometries were originally obtained in Ref. [214], they were recalculated in Ref. [183] as part of the work done for this dissertation.

<sup>b</sup>Taken from Ref. [219].



The MP2 and CCSD(T) geometries optimized in this study, along with selected DFT results for the  $S_1$  isomer taken from Ref. [219], are presented in Table 5.4. For the MP2 results, the size of the basis has usually a slight effect on the calculated geometries, with the majority of the differences between the MP2/SBKJC(1f) and MP2/cc-pVDZ-PP optimizations falling into the  $\sim 0.01$  Å range, and with the largest differences being 0.031 and 0.081 Å. Similar remarks apply to the CCSD(T) geometries obtained using the two different basis sets exploited here, with the average and maximum differences being 0.014 and 0.056 Å, respectively. Comparing the MP2 and CCSD(T) geometries within a given basis, we see that for the SBKJC(1f) results, the differences range from 0.051 to 0.129 Å, for 0.069 Å on average, while for the cc-pVDZ-PP basis they are as large as 0.154 Å, again 0.069 Å on average. Generally, the CCSD(T) calculations make the Au–Au bonds in  $\text{Au}_8$  longer. This might be a consequence of MP2 overbinding the  $\text{Au}_8$  cluster, which has some characteristics of the weakly bound systems, including the potentially important role of non-additive dispersion forces [205, 420–422]. Indeed, the binding energy per gold atom obtained in the MP2/cc-pVTZ-PP+CV//MP2/cc-pVDZ-PP calculations, averaged over the four structures of  $\text{Au}_8$  examined in this work, is 57.43 kcal/mol, as opposed to 45.26 kcal/mol obtained in the corresponding CCSD(T)/cc-pVTZ-PP+CV//MP2/cc-pVDZ-PP computations (45.16 kcal/mol when the CCSD(T)/cc-pVDZ-PP geometries are used). However, as shown below, these differences between the MP2 and CCSD(T) geometries, although significant in some cases, are not sufficiently large to affect the final conclusions regarding the planarity *vs* non-planarity of the  $\text{Au}_8$  particle. What seems to be more important is the electron correlation treatment applied to the relative energies of the various structures.

The relative energetics, shown in Table 5.5, demonstrate that even when the geometries are optimized with a larger cc-pVDZ-PP basis set and the CV correlations are included in the calculations, the MP2 results are completely unreliable. The use of the cc-pVTZ-PP basis set, with or without CV correlations, in MP2 calculations does not help either. The MP2 calculations arrange the  $S_6$  and  $S_1$  isomers as the lowest and highest in energy, respectively, while strongly favoring the non-planar  $Au_8$  structures over the planar  $S_1$  isomer. This immediately implies that one needs to account for the higher-order correlation effects to obtain reliable energetics of the various  $Au_8$  structures.

The CCSD(T)/cc-pV $x$ Z-PP and CCSD(T)/cc-pV $x$ Z-PP+CV ( $x = D, T$ ) single-point calculations, using the MP2/cc-pVDZ-PP geometries, lower the energy differences of the  $S_3$ ,  $S_4$ , and  $S_6$  structures, relative to  $S_1$ , by approximately 1–2 kcal/mol compared to their counterparts using the MP2/SBKJC(1f) geometries reported in Ref. [217], but they do not alter the main conclusions of Ref. [217] (or Refs. [215] and [216]) that  $S_1$  is a global minimum. According to our best CCSD(T)/cc-pVTZ-PP+CV calculations employing the geometries optimized at the CCSD(T)/cc-pVDZ-PP level, the energy ordering of the lowest four isomers of  $Au_8$  is  $S_1 < S_3 \approx S_4 < S_6$ , in agreement with the findings of Ref. [217]. However, it is important to emphasize that the use of the smaller cc-pVDZ-PP basis set in the CCSD(T) calculations, although reasonable in geometry optimizations, is not sufficient to provide a reliable description of the relative energies of the various  $Au_8$  structures.

Indeed, as shown in Table 5.5, the use of the cc-pVDZ-PP basis set in the CCSD(T) calculations instead of cc-pVTZ-PP lowers the energies of the  $S_3$ ,  $S_4$ , and  $S_6$  structures relative to  $S_1$  by 4–5 kcal/mol, even when the CCSD(T)/cc-pVDZ-PP geometries are employed,

bringing  $S_3$ ,  $S_4$ , and  $S_6$  close to  $S_1$  when the CV effects are ignored, and placing  $S_3$ ,  $S_4$ , and  $S_6$  below  $S_1$  when the CV effects are accounted for. One has to use the basis set of the cc-pVTZ quality in the CCSD(T) calculations to stabilize  $S_1$  as a global minimum. The CV effects play some role, lowering the energies of  $S_3$ ,  $S_4$ , and  $S_6$  relative to  $S_1$  by about 1–2 kcal/mol at the CCSD(T)/cc-pVTZ-PP level, but they do not alter the relative energies when the cc-pVTZ-PP basis set is used in the CCSD(T) calculations (note that they do alter them when the smaller cc-pVDZ-PP basis set is employed). This is very different from the our study of the  $\text{Au}_3^-$  photoelectron spectrum [184] where the CV correlations had a significant effect on the final results, which demonstrates the need to carefully consider the system one wishes to study as to how it needs to be treated in order to obtain an accurate description of such. Similar applies to the role of geometry optimization. The use of the better CCSD(T)/cc-pVDZ-PP geometries instead of the MP2/SBKJC(1f) geometries employed in Ref. [217] lowers the energies of the  $S_3$ ,  $S_4$ , and  $S_6$  structures relative to  $S_1$  by 1–2 kcal/mol, but this is not sufficient to change the overall energy ordering, as long as the cc-pVTZ-PP basis set is used in the final single-point CCSD(T) calculations.

In summary, we performed geometry optimizations at the MP2/cc-pVDZ-PP, CCSD(T)/SBKJC(1f), and CCSD(T)/cc-pVDZ-PP levels of theory employing relativistic ECPs for the four  $\text{Au}_8$  lowest-energy isomers considered in the earlier CC studies [214–217]. We then used the resulting geometries to carry out the single-point MP2/cc-pVDZ-PP, MP2/cc-pVTZ-PP, CCSD(T)/cc-pVDZ-PP, and CCSD(T)/cc-pVTZ-PP energy calculations, adopting the same ECPs as used in the geometry optimizations and correlating the  $5d^{10}6s^1$  valence and  $5s^25p^65d^{10}6s^1$  semi-core and valence electrons. According to our best CCSD(T) calcula-

Table 5.5: Relative energies (in kcal/mol) of the S<sub>1</sub>, S<sub>3</sub>, S<sub>4</sub>, and S<sub>6</sub> isomers of Au<sub>8</sub>, with respect to the S<sub>1</sub> isomer, obtained in the MP2 and CCSD(T) calculations.

| Method/Basis                    |                                    | Isomer Energy  |                |                |                |
|---------------------------------|------------------------------------|----------------|----------------|----------------|----------------|
| Geometry Optimization           | Single-point Calculation           | S <sub>1</sub> | S <sub>3</sub> | S <sub>4</sub> | S <sub>6</sub> |
| MP2/SBKJC(1f)                   | MP2/SBKJC(1f) <sup>a,b</sup>       | 0.0            | -25.1          | -23.4          | -30.8          |
|                                 | MP2/cc-pVDZ-PP <sup>b</sup>        | 0.0            | -16.9          | -15.8          | -22.3          |
|                                 | MP2/cc-pVDZ-PP+CV <sup>b</sup>     | 0.0            | -19.9          | -19.2          | -25.5          |
|                                 | MP2/cc-pVTZ-PP <sup>b</sup>        | 0.0            | -15.4          | -16.3          | -21.8          |
|                                 | MP2/cc-pVTZ-PP+CV <sup>b</sup>     | 0.0            | -16.5          | -17.8          | -23.3          |
|                                 | CCSD(T)/SBKJC(1f) <sup>a,b</sup>   | 0.0            | -4.7           | -2.5           | -3.2           |
|                                 | CCSD(T)/cc-pVDZ-PP <sup>b</sup>    | 0.0            | 3.6            | 5.2            | 5.6            |
|                                 | CCSD(T)/cc-pVDZ-PP+CV <sup>b</sup> | 0.0            | 0.1            | 1.3            | 1.8            |
|                                 | CCSD(T)/cc-pVTZ-PP <sup>b</sup>    | 0.0            | 7.5            | 7.5            | 9.1            |
|                                 | CCSD(T)/cc-pVTZ-PP+CV <sup>b</sup> | 0.0            | 6.0            | 5.7            | 7.2            |
| CCSD(T)/SBKJC(1f) <sup>c</sup>  | CCSD(T)/SBKJC(1f)                  | 0.0            | -6.29          | -4.04          | -5.08          |
|                                 | CCSD(T)/cc-pVDZ-PP                 | 0.0            | 0.37           | 1.99           | 1.98           |
|                                 | CCSD(T)/cc-pVDZ-PP+CV              | 0.0            | -3.05          | -1.72          | -1.74          |
|                                 | CCSD(T)/cc-pVTZ-PP                 | 0.0            | 5.19           | 5.41           | 6.49           |
|                                 | CCSD(T)/cc-pVTZ-PP+CV              | 0.0            | 3.93           | 3.85           | 4.86           |
| MP2/cc-pVDZ-PP <sup>c</sup>     | MP2/cc-pVDZ-PP                     | 0.0            | -17.32         | -16.37         | -22.52         |
|                                 | MP2/cc-pVDZ-PP+CV                  | 0.0            | -19.81         | -19.31         | -25.24         |
|                                 | MP2/cc-pVTZ-PP                     | 0.0            | -14.39         | -15.25         | -20.95         |
|                                 | MP2/cc-pVTZ-PP+CV                  | 0.0            | -15.04         | -16.33         | -21.92         |
|                                 | CCSD(T)/cc-pVDZ-PP                 | 0.0            | 1.47           | 3.04           | 3.41           |
|                                 | CCSD(T)/cc-pVDZ-PP+CV              | 0.0            | -1.81          | -0.56          | -0.11          |
|                                 | CCSD(T)/cc-pVTZ-PP                 | 0.0            | 6.52           | 6.54           | 7.83           |
|                                 | CCSD(T)/cc-pVTZ-PP+CV              | 0.0            | 5.33           | 5.00           | 6.28           |
| CCSD(T)/cc-pVDZ-PP <sup>c</sup> | CCSD(T)/cc-pVDZ-PP                 | 0.0            | 0.16           | 1.72           | 1.79           |
|                                 | CCSD(T)/cc-pVDZ-PP+CV              | 0.0            | -3.08          | -1.74          | -1.72          |
|                                 | CCSD(T)/cc-pVTZ-PP                 | 0.0            | 5.63           | 5.72           | 6.91           |
|                                 | CCSD(T)/cc-pVTZ-PP+CV              | 0.0            | 4.60           | 4.39           | 5.57           |

<sup>a</sup>From Ref. [214].

<sup>c</sup>From Ref. [217].

<sup>c</sup>From Ref. [183] done as part of the work for this dissertation.

tions with the cc-pVTZ-PP basis set, accounting for the CV correlations, and using the CCSD(T)/cc-pVDZ-PP optimized geometries, the lowest-energy structure of Au<sub>8</sub> is the planar S<sub>1</sub> isomer shown in Fig. 5.1. This is in agreement with an experimental study [423] published after our initial study was published in which they showed, using Far-IR spectroscopy, that the lowest energy isomer of Au<sub>8</sub> is indeed the planar S<sub>1</sub> structure. The next two isomers, S<sub>3</sub> and S<sub>4</sub>, are nearly degenerate and lie about 4–5 kcal/mol above S<sub>1</sub>, and the fourth structure studied in this work, designated as S<sub>6</sub>, is about 6 kcal/mol above S<sub>1</sub>. Geometry relaxation plays a rather small role on the calculated energies, lowering the S<sub>3</sub>, S<sub>4</sub>, and S<sub>6</sub> structures relative to S<sub>1</sub> by 1–2 kcal/mol, but does not alter the main conclusions about the planarity of the lowest-energy isomer of Au<sub>8</sub>, as long as one uses the CCSD(T) approach and the larger basis set of the triple-zeta quality, such as the cc-pVTZ-PP basis employed in this dissertation. At the same time, use of the MP2 method provides completely incorrect energetic results, even when the larger basis sets and the CV correlations are included, in agreement with the findings of Ref. [217]. Thus, one needs to use high-level methods, such as CCSD(T), capable of accounting for higher-order correlation effects, to obtain a reliable description at the *ab initio* wave function theory level. This may be related to the fact that gold particles are characterized by substantial non-additive dispersion effects, which cannot be captured by MP2.

## 5.2 Unrestricted implementation of CR-CC(2,3)

The success of CC theory is typically associated with the popular single reference (SR) CCSD(T) approach, which provides an accurate description of dynamical electron correla-

tion effects for nondegenerate ground states of molecules near the equilibrium geometries (cf. Section 5.1.3). It is well known that the RHF-based CCSD(T), RCCSD(T), approximation fails when applied to biradicals and bond breaking situations where nondynamical correlation effects become more significant. It is also well known that one can remedy the unphysical characterization manifest in the potential energy surface (PES) along bond breaking coordinates of single bond breaking situations in the RCCSD(T) computations by switching to the UHF-based CCSD(T), UCCSD(T), approach. While the overall energetics of the PES along bond breaking coordinates are improved it is done at the sacrifice of introducing other unphysical characteristics, such as spin contamination and non-analytic behavior of the PES.

The CR-CC methods, such as the original CR-CCSD(T) [40, 41, 87–89] and newer rigorously size-extensive left eigenvalue approaches, CR-CC(2,3), as well as their higher-order extensions [40, 41, 87–89, 93, 94, 424], have been shown to improve the poor RCCSD(T) results, particularly in regions of the PES involving single bond breaking even when a RHF determinant is used as the reference wave function (see Ref. [68] and references therein for further details). The CR-CCSD(T) and CR-CC(2,3) approaches are of particular interest as they retain the same black-box ease of use and iterative  $\mathcal{N}^6$  and noniterative  $\mathcal{N}^7$  CPU scaling costs as CCSD(T). While the performance of the CR-CCSD(T) and CR-CC(2,3) approximations employing an RHF wave function, RCR-CCSD(T) and RCR-CC(2,3), respectively, has been extensively examined, it is not well known if using an UHF reference determinant will offer some improvement of the already quite good CR-CC PESs along bond breaking coordinates for closed-shell singlet molecular systems. Recently the UHF-based CR-CCSD(T), UCR-CCSD(T), approach was used to examine single bond breaking on a singlet PES [425],

concluding that, while UCR-CCSD(T) computations give qualitative correct results, the results at stretched intermediate bond lengths over estimate the relative energies, the same as seen with UCCSD(T) calculations. This behavior is a known feature of spin contaminated UHF-based CC results [426].

As part of this thesis, we examined whether employing an UHF reference wave function in the CR-CC calculations of single bond breaking into open-shell fragments on singlet PESs will offer any improvements over the already accurate RHF-based CR-CC results by examining bond breaking in the two model diatomic systems used for benchmarking, namely, HF and F<sub>2</sub> as well as two polyatomic systems, which serve as models for larger molecules where important O–O and C–C single bond breaking events can play crucial roles, the H<sub>2</sub>O<sub>2</sub>, and C<sub>2</sub>H<sub>6</sub> molecular systems, comparing them with the exact, full configuration interaction (CI), and full CCSDT results. We also compare our UHF-based CR-CC results with their RHF-based counterparts showing that the spin-adapted RHF-based CR-CC(2,3) results provide the most accurate description at all points along the PESs.

### 5.2.1 Algorithm

In order to compute the UHF-based CR-CC(2,3) results we needed to transform the integrals in the atomic orbital (AO) basis to the molecular orbital (MO) basis. This is done using two- and four-index integral transformations, with the latter being the more expensive step. The four-index integral transformation of the two-body integrals can be represented algebraically as

$$(pq|rs) = \sum_{\mu,\nu,\rho,\sigma} C_{p\mu}C_{q\nu}C_{r\rho}C_{s\sigma}(\mu\nu|\rho\sigma), \quad (5.17)$$

where  $(pq|rs)$  and  $(\mu\nu|\rho\sigma)$  are the two-electron integrals in the MO and AO bases, respectively, and the  $C$  coefficients are the standard expansion coefficients obtained from diagonalizing the Fock matrix. We only considered the  $\alpha$ - $\alpha$ ,  $\beta$ - $\beta$ , and  $\alpha$ - $\beta$  blocks of the two-electron integrals as other combinations are zero. If implemented naively, looping over all  $p, q, r, s, \mu, \nu, \rho,$  and  $\sigma$  using nested loops this would be an  $O(N^8)$  algorithm and would not be very useful. In developing our own routine we made use of the fact that modern Fortran allows for the use of up to 8-dimensional arrays, storing the two-body matrix elements in a 4-dimensional array. We would then take a slice of the four dimensional array and perform the appropriate matrix-matrix multiplication using calls to the BLAS DGEMM routine. We also made use of the threading capabilities of the Intel MKL library, allowing for a level of parallelism in our routine. The algorithm is trivially parallelizable in the way that it is written and will be pursued in future studies. One criticism of this type of approach is that, depending on the order of calls, it might result in a large number of cache misses as compared to a routine where the 4-dimensional array is flattened to two or one dimensions, which for large problems would result in a drop in performance. Even for a properly optimized program making use of the 4-dimensional array there might still be a loss in efficiency compared to one that flattens the array due to the contiguous nature of current CPU memory, particularly when run in serial. However, as mentioned above, it is written such that it is an “embarrassingly parallel” algorithm making it easy to parallelize. Once the one- and two-body integrals were transformed to the MO basis they were read into our group’s in-house CC program package, called CC\_PACKAGE, which we used to compute the UHF-based CR-CC(2,3) results discussed in the next section.



### 5.2.2 Numerical examples

We used the NWChem suite to carry out the UHF-based CR-CCSD(T) calculations and the GAMESS package for the RHF-based CCSD(T), CR-CCSD(T), and CR-CC(2,3) computations, which spin-adapted routines were implemented by our group. The RHF- and UHF-based CCSDT calculations reported here were done using in-house spin-integrated CC codes. To perform the UHF-based CR-CC(2,3) calculations we used a modified version of GAMESS to obtain the proper one- and two-body integrals and coefficients, which were then transformed with our own four-index integral transformation routine, as discussed above, and finally carried out the UHF-based CR-CC(2,3) computations using our in-house CC routines, which work for any SR wave function. The specific details of basis sets and number of frozen core orbitals are discussed below with the results of each example.

#### 5.2.2.1 The HF molecule

Following previous CC studies [87, 88, 90, 93] we computed the PES of the HF molecule as described by the DZ basis set [185] and correlating all electrons, for which the full CI energies [361] are available for comparison. The results of our RHF- and UHF-based CC and CR-CC calculations for HF are presented in Tables 5.6 and 5.7 and graphically in Figs. 5.2(a) and 5.3(a). As mentioned above and as is well known the UHF-based CCSD(T) results are a dramatic improvement over their RHF-based counterparts, reducing the mean signed error (MSE) from around -20.0 millihartree to 0.5–0.7 millihartree when compared to RCCSDT and UCCSDT, respectively, with the same degree of improvement compared to full CI. The mean unsigned error (MUE) for the RHF and UHF-based CCSD(T) calculations reduces

from  $\sim 7$  to 0.5–0.7 millihartree compared to RCCSDT and UCCSDT, respectively. The nonparallelity error (NPE) shows a massive decrease when the UHF reference determinant is employed in the CCSD(T) computations, dropping from about 50 millihartree to around 1 millihartree compared to their respective parent CCSDT approaches and full CI.

Examining the CR-CC results, the already reasonable MSE, MUE and NPE values for the RHF-based approaches are improved upon slightly when the UHF reference determinant is employed. Our least accurate CR-CC approach, which still provides very good results, CR-CCSD(T) provides reasonable MSE and MUE values on the order of 1 (1.5) millihartree compared to CCSDT (full CI) when the RHF wave function is used as a reference and an NPE of about 1.4 (1.6) millihartree compared to full CCSDT (full CI). Switching to the UHF wave function as a reference the MSE and MUE values improve by about 0.01 (0.4) millihartree compared to UCCSDT (full CI). In this case, though, we see a slight increase in the NPE going from 1.43 (1.60) to 1.85 (2.58) millihartree compared to RCCSDT and UCCSDT (full CI), respectively. This is a manifestation of the well known spin-contamination problem in the CC calculations [426] and, as we will show in this letter, is present in all of the approximate triples CC approaches examined in this work when the UHF reference wave function is employed. The CR-CC(2,3),A [CCSD(2)<sub>T</sub>] approach shows slightly better performance with the RHF-based CR-CC(2,3),A MSE and MUE values being  $\sim 0.9$  (1.3) millihartree compared to RCCSDT (full CI) and NPE on the order of 1.5 (2.0) millihartree. The UHF-based CR-CC(2,3),A show a similar behavior as with its UHF-based counterpart CR-CCSD(T), lowering the MSE and MUE by 0.01 millihartree and the NPE increasing from around 1.5 (2.0) to 1.7 (2.4) compared to RCCSDT and UCCSDT (full CI), respectively. Our best, CR-

CC(2,3),D, computations provide the best overall agreement when compared to their parent CCSDT results and the exact, full CI, values. The RHF-based CR-CC(2,3),D MSE and MUE -0.89 (-0.29) and 0.37 (0.27) millihartree are in very good agreement with the RHF-based CCSDT (full CI) with the UHF-based CR-CC(2,3),D offer some improvement, particularly for the MSE, changing its value to -0.04 (0.18) millihartree with respect to UCCSDT (full CI). The increase in the NPE when using the RHF *vs* UHF reference determinant is not as severe as in the CR-CCSD(T) and CR-CC(2,3),A cases due in part to the more robust treatment using the full expression of the denominator for the CR-CC(2,3),D calculations.

### 5.2.2.2 The F<sub>2</sub> molecule

The very challenging F<sub>2</sub> molecule, long known to have a large degree of nondynamical correlation effects and to be unbound when described at the UHF level of theory, is ideal for testing approximate triples CC approaches as it requires a proper balance of the dynamical and nondynamical correlation effects for an accurate description of its PES. It is also well known that CCSDT can accurately describe single bond breaking situations, as the inclusion of the effects due to triply excited clusters are required to properly account for the important correlations in single bond breaking. Thus F<sub>2</sub> has long provided a way to test the reliability and robustness of approximate triples CC methodologies and their ability to balance dynamical and nondynamical correlations. In the calculations reported here on F<sub>2</sub> we used the spherical components of *d* basis functions of the cc-pVDZ basis set [186], freezing the core 1s orbitals of each F atom. The results at a few F–F distances and the statistics for all the F<sub>2</sub> results shown in Tables 5.8 and 5.9, and Figs. 5.2(b) (compared to full CI) and 5.3(b) (compared to CCSDT).

Table 5.6: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with the corresponding full CI data<sup>a</sup> for the equilibrium and four displaced geometries ( $R_e = 1.7328$  bohr) of the HF molecule, as described by the spherical Dunning DZ basis.

| Method                   |     | $R_e^a$     | $2R_e^a$    | $3R_e^a$   | $5R_e^a$   | MSE <sup>b</sup> | MUE <sup>c</sup> | NPE <sup>d</sup> |
|--------------------------|-----|-------------|-------------|------------|------------|------------------|------------------|------------------|
| CCSD                     | RHF | 1.63        | 6.05        | 11.60      | 12.29      | 7.89             | 7.89             | 10.66            |
|                          | UHF | 1.63        | 5.89        | 2.47       | 1.23       | 2.80             | 2.80             | 4.66             |
| CC(T) <sup>e</sup>       | RHF | 0.33        | 0.04        | -24.48     | -53.18     | -19.32           | 7.27             | 53.51            |
|                          | UHF | 0.33        | 1.26        | 1.22       | 0.04       | 0.71             | 0.71             | 1.22             |
| CR(T) <sup>e</sup>       | RHF | 0.50        | 2.03        | 2.10       | 1.65       | 1.57             | 1.57             | 1.60             |
|                          | UHF | 0.50        | 2.74        | 1.34       | 0.15       | 1.18             | 1.18             | 2.58             |
| CR(2,3),A <sup>e,f</sup> | RHF | 0.23        | 1.45        | 2.18       | 1.44       | 1.32             | 1.32             | 1.95             |
|                          | UHF | 0.23        | 2.40        | 1.18       | -0.01      | 0.95             | 0.95             | 2.41             |
| CR(2,3),D <sup>e</sup>   | RHF | -0.12       | 0.06        | -0.10      | -1.00      | -0.29            | 0.27             | 1.07             |
|                          | UHF | -0.12       | 0.36        | 0.83       | -0.34      | 0.18             | 0.41             | 1.17             |
| T <sup>e</sup>           | RHF | 0.17        | 0.86        | 0.96       | 0.43       | 0.60             | 0.60             | 0.78             |
|                          | UHF | 0.17        | 0.63        | 0.18       | -0.10      | 0.22             | 0.27             | 0.73             |
| Full CI <sup>g</sup>     |     | -100.160300 | -100.021733 | -99.985281 | -99.983293 |                  |                  |                  |

<sup>a</sup>Full CI results were taken from Ref. [361].

<sup>b</sup>Mean signed error relative to full CI.

<sup>c</sup>Mean unsigned error relative to full CI.

<sup>d</sup>Nonparallelity error relative to full CI.

<sup>e</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to full CI. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>f</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>g</sup>The total full CI energies in hartree.

Table 5.7: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach data for the equilibrium and four displaced geometries ( $R_e = 1.7328$  bohr) of the HF molecule, as described by the spherical Dunning DZ basis.

| Method                   |     | $R_e^a$     | $2R_e^a$    | $3R_e^a$   | $5R_e^a$   | MSE <sup>a</sup> | MUE <sup>b</sup> | NPE <sup>c</sup> |
|--------------------------|-----|-------------|-------------|------------|------------|------------------|------------------|------------------|
| CCSD                     | RHF | 1.46        | 5.19        | 10.64      | 11.86      | 7.29             | 7.29             | 10.40            |
|                          | UHF | 1.46        | 5.26        | 2.29       | 1.33       | 2.59             | 2.59             | 3.93             |
| CC(T) <sup>d</sup>       | RHF | 0.15        | -0.82       | -25.44     | -53.61     | -19.93           | 7.29             | 53.77            |
|                          | UHF | 0.15        | 0.63        | 1.04       | 0.14       | 0.49             | 0.49             | 0.90             |
| CR(T) <sup>d</sup>       | RHF | 0.33        | 1.18        | 1.14       | 1.22       | 0.97             | 0.97             | 0.89             |
|                          | UHF | 0.33        | 2.11        | 1.17       | 0.26       | 0.96             | 0.96             | 1.85             |
| CR(2,3),A <sup>d,e</sup> | RHF | 0.06        | 0.59        | 1.22       | 1.01       | 0.72             | 0.72             | 1.16             |
|                          | UHF | 0.06        | 1.77        | 1.00       | 0.09       | 0.73             | 0.73             | 1.71             |
| CR(2,3),D <sup>d</sup>   | RHF | -0.29       | -0.79       | -1.05      | -1.44      | -0.89            | 0.37             | 0.76             |
|                          | UHF | -0.29       | -0.27       | 0.65       | -0.24      | -0.04            | 0.36             | 0.94             |
| T <sup>d,f</sup>         | RHF | -100.160127 | -100.020878 | -99.984324 | -99.982862 |                  |                  |                  |
|                          | UHF | -100.160127 | -100.021105 | -99.985103 | -99.983395 |                  |                  |                  |

<sup>a</sup>Mean signed error relative to CCSDT.

<sup>b</sup>Mean unsigned error relative to CCSDT.

<sup>c</sup>Nonparallelity error relative to CCSDT.

<sup>d</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to the respective RHF- and UHF-based full CCSDT. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>e</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>f</sup>The total CCSDT energies in hartree.

Examining the RHF- and UHF-based CCSD(T) results we see the well known behavior, with the RHF-based CCSD(T) calculations providing unphysical results at larger internuclear distances, resulting in an MSE of about -7 (-5) millihartree, MUE of  $\sim 7$  ( $\sim 7$ ) millihartree, and a very large NPE value of 40 (41) millihartree with respect to the RHF-based CCSDT (full CI) approach. The CCSD(T) results using the UHF determinant as a reference wave function improve the RHF-based CCSD(T) computations, changing the MSE and MUE both to about 3 (4) millihartree, with the most dramatic change in the NPE to the value of approximately 10 (11) millihartree compared to the parent full UCCSDT (full CI) method, demonstrating once again the strong dependence of the CCSD(T) results on the reference determinant employed.

Again, we can see the quality of the CR-CC results are not plagued by this reliance on the type of determinant used as reference, with the CR-CCSD(T) and CR-CC(2,3),A approaches providing very similar results, with the CR-CC(2,3),A variant improving the CR-CCSD(T) calculations by about 1 millihartree as measured by the MSE, MUE, and NPE values when the RHF wave function is used and by about 0.6 millihartree for the UHF-based results, compared to both CCSDT and the full CI energies. If we compare their RHF- and UHF-based computations we again see a slight improvement in their MSE and MUE values of less than 1 millihartree. Examining the NPE values for the RHF-based and CR-CCSD(T) and CR-CC(2,3),A approaches, which are on the order of about 7 millihartree, we see that switching to their UHF-based counterparts, the NPEs increase by 4–5 millihartree due to spin contamination of the results at the intermediate internuclear bond distances. Examining our best CR-CC(2,3),D results, we see excellent agreement of our RHF-based results with

CCSDT (full CI), providing MSE of about 0.6 (2.3), MUE of approximately 0.9 (2.3), and NPE of  $\sim 2.1$  (4.0) millihartree. When we employ the UHF determinant the overall statistics of our CR-CC(2,3),D, while improving on the UCCSD(T) results, are slightly worse than the RHF-based CR-CC(2,3),D calculations, with MSE and MUE of about 2 (3) millihartree with respect to CCSDT (full CI) and again a much larger NPE of  $\sim 9$  ( $\sim 10$ ) millihartree, which is still an improvement of the UCCSD(T) calculations.

As shown in Fig. 5.3(b) all of the UHF-based approximate triples approaches have large errors from their parent CCSDT method with the CR-CC(2,3),D variant giving the smallest error of 8.5 millihartree (5.4 kcal/mol) in this intermediate stretched bond length region. In this same region the RHF-based CR-CC(2,3),D approach provides errors of about 1 millihartree ( $< 1$  kcal/mol). While the UHF-based CR-CC and CCSD(T) computations approach the full CCSDT results in the asymptotic region the large errors similar in size to some barriers for chemical reactions make their use problematic. The RHF-based CR-CC(2,3),D, on the other hand, provides a reasonable description of the PES at all internuclear distances with the errors in the asymptotic region being  $\sim 1$  kcal/mol, typically called chemical accuracy. Another important feature of the robust RHF-based CR-CC(2,3),D approach, as demonstrated in Fig. 5.2(b), which shows the errors relative to full CI, is that at all points along the PES its computed energies follow very closely the full CCSDT values, being shifted by an approximately equal amount, demonstrating the systematically improvable nature and robustness of this approach.

Table 5.8: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with the corresponding full CI data for the equilibrium and displaced geometries of the F<sub>2</sub> molecule, as described by the spherical cc-pVDZ basis set.

| Method                   |     | 1.14 | 1.20  | 1.30  | 1.36  | $R_e^a$ | 1.50  | 1.60  | 1.80  | 2.00  | 2.20  | 2.40   | 2.80   | 8.00   | MSE <sup>b</sup> | MUE <sup>c</sup> | NPE <sup>d</sup> |
|--------------------------|-----|------|-------|-------|-------|---------|-------|-------|-------|-------|-------|--------|--------|--------|------------------|------------------|------------------|
| CCSD                     | RHF | 5.81 | 6.61  | 8.37  | 9.58  | 10.74   | 13.14 | 16.27 | 23.56 | 30.76 | 37.10 | 41.97  | 46.92  | 50.62  | 23.19            | 23.19            | 44.81            |
|                          | UHF | 5.81 | 6.61  | 8.37  | 9.98  | 11.96   | 15.72 | 19.54 | 20.87 | 14.15 | 7.98  | 4.73   | 2.42   | 2.07   | 10.02            | 10.02            | 18.80            |
| CC(T) <sup>e</sup>       | RHF | 0.88 | 1.03  | 1.40  | 1.59  | 1.72    | 2.04  | 2.27  | 1.89  | -0.71 | -5.18 | -10.66 | -21.21 | -38.71 | -4.90            | 6.87             | 40.98            |
|                          | UHF | 0.88 | 1.03  | 1.40  | 2.00  | 2.79    | 4.12  | 6.15  | 11.78 | 10.23 | 5.67  | 2.81   | 0.63   | 0.30   | 3.83             | 3.83             | 11.48            |
| CR(T) <sup>e</sup>       | RHF | 1.58 | 1.84  | 2.48  | 2.88  | 3.24    | 4.07  | 5.12  | 7.41  | 9.10  | 10.29 | 10.91  | 10.42  | 7.18   | 5.89             | 5.89             | 9.33             |
|                          | UHF | 1.58 | 1.84  | 2.48  | 3.48  | 4.91    | 7.39  | 10.06 | 13.68 | 10.79 | 5.96  | 3.04   | 0.84   | 0.50   | 5.12             | 5.12             | 13.17            |
| CR(2,3),A <sup>e,f</sup> | RHF | 1.32 | 1.56  | 2.16  | 2.53  | 2.85    | 3.62  | 4.58  | 6.60  | 7.95  | 8.76  | 9.06   | 8.18   | 4.78   | 4.92             | 4.92             | 7.74             |
|                          | UHF | 1.32 | 1.56  | 2.16  | 2.79  | 3.69    | 5.68  | 8.38  | 13.05 | 10.59 | 5.80  | 2.89   | 0.70   | 0.37   | 4.54             | 4.54             | 12.68            |
| CR(2,3),D <sup>e</sup>   | RHF | 0.42 | 0.53  | 0.87  | 1.06  | 1.20    | 1.67  | 2.22  | 3.36  | 3.96  | 4.30  | 4.37   | 3.61   | 2.62   | 2.32             | 2.32             | 3.95             |
|                          | UHF | 0.43 | 0.53  | 0.86  | 1.27  | 1.83    | 3.15  | 5.19  | 10.52 | 9.69  | 5.45  | 2.67   | 0.46   | -0.07  | 3.23             | 3.24             | 10.59            |
| T <sup>e</sup>           | RHF | 0.75 | 0.86  | 1.19  | 1.34  | 1.45    | 1.79  | 2.08  | 2.59  | 2.52  | 2.42  | 2.38   | 1.74   | 1.02   | 1.70             | 1.70             | 1.84             |
|                          | UHF | 0.75 | 0.86  | 1.19  | 1.40  | 1.57    | 1.89  | 2.06  | 1.97  | 1.37  | 1.04  | 0.99   | 0.36   | 0.15   | 1.20             | 1.20             | 1.91             |
| Full CI <sup>g</sup>     |     | 7.18 | 48.11 | 85.10 | 95.17 | 99.20   | 99.81 | 95.10 | 80.90 | 68.82 | 61.65 | 58.23  | 55.77  | 55.45  |                  |                  |                  |

<sup>a</sup>Experimental equilibrium F–F internuclear bond length  $R_e = 1.41193 \text{ \AA}$  taken from Ref. [427].

<sup>b</sup>Mean signed error relative to full CI.

<sup>c</sup>Mean unsigned error relative to full CI.

<sup>d</sup>Nonparallelity error relative to full CI.

<sup>e</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to full CI. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>f</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>g</sup>The total CEEIS full CI energies  $E$ , reported as  $-(199 + mE)$ , are in hartree, taken from Ref. [428].



Table 5.9: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the F<sub>2</sub> molecule, as described by the spherical cc-pVDZ basis set.

| Method                   |     | 1.14  | 1.20  | 1.30  | 1.36  | $R_e^a$ | 1.50  | 1.60  | 1.80  | 2.00  | 2.20  | 2.40   | 2.80   | 8.00   | MSE <sup>b</sup> | MUE <sup>c</sup> | NPE <sup>d</sup> |
|--------------------------|-----|-------|-------|-------|-------|---------|-------|-------|-------|-------|-------|--------|--------|--------|------------------|------------------|------------------|
| CCSD                     | RHF | 5.06  | 5.74  | 7.18  | 8.24  | 9.29    | 11.37 | 14.19 | 20.97 | 28.24 | 34.68 | 39.59  | 45.18  | 49.61  | 21.49            | 21.49            | 44.54            |
|                          | UHF | 5.06  | 5.74  | 7.18  | 8.59  | 10.39   | 13.83 | 17.48 | 18.91 | 12.78 | 6.94  | 3.74   | 2.07   | 1.92   | 8.82             | 8.82             | 16.84            |
| CC(T) <sup>e</sup>       | RHF | 0.13  | 0.16  | 0.22  | 0.25  | 0.27    | 0.28  | 0.18  | -0.70 | -3.23 | -7.61 | -13.05 | -22.95 | -39.72 | -6.60            | 6.68             | 40.00            |
|                          | UHF | 0.13  | 0.16  | 0.22  | 0.60  | 1.22    | 2.23  | 4.10  | 9.82  | 8.87  | 4.63  | 1.82   | 0.28   | 0.16   | 2.63             | 2.63             | 9.68             |
| CR(T) <sup>e</sup>       | RHF | 0.84  | 0.98  | 1.30  | 1.54  | 1.79    | 2.31  | 3.04  | 4.82  | 6.59  | 7.87  | 8.53   | 8.68   | 6.17   | 4.19             | 4.19             | 7.85             |
|                          | UHF | 0.84  | 0.98  | 1.30  | 2.09  | 3.34    | 5.50  | 8.01  | 11.71 | 9.42  | 4.91  | 2.05   | 0.48   | 0.36   | 3.92             | 3.92             | 11.36            |
| CR(2,3),A <sup>e,f</sup> | RHF | 0.58  | 0.70  | 0.97  | 1.19  | 1.41    | 1.86  | 2.49  | 4.01  | 5.43  | 6.33  | 6.67   | 6.44   | 3.76   | 3.22             | 3.22             | 6.10             |
|                          | UHF | 0.58  | 0.70  | 0.97  | 1.40  | 2.12    | 3.79  | 6.33  | 11.08 | 9.22  | 4.76  | 1.90   | 0.34   | 0.22   | 3.34             | 3.34             | 10.86            |
| CR(2,3),D <sup>e</sup>   | RHF | -0.32 | -0.33 | -0.32 | -0.28 | -0.25   | -0.10 | 0.13  | 0.77  | 1.45  | 1.87  | 1.99   | 1.87   | 1.61   | 0.62             | 0.87             | 2.08             |
|                          | UHF | -0.31 | -0.33 | -0.32 | -0.13 | 0.26    | 1.26  | 3.13  | 8.55  | 8.32  | 4.41  | 1.69   | 0.10   | -0.22  | 2.03             | 2.23             | 8.88             |
| T <sup>g</sup>           | RHF | 6.44  | 47.25 | 83.91 | 93.83 | 97.75   | 98.05 | 93.02 | 78.31 | 66.30 | 59.23 | 55.85  | 54.03  | 54.43  |                  |                  |                  |
|                          | UHF | 6.44  | 47.25 | 83.91 | 93.77 | 97.63   | 97.92 | 93.05 | 78.93 | 67.45 | 60.61 | 57.24  | 55.41  | 55.30  |                  |                  |                  |

<sup>a</sup>Experimental equilibrium F–F internuclear bond length  $R_e = 1.41193$  Å taken from Ref. [427].

<sup>b</sup>Mean signed error relative to CCSDT.

<sup>c</sup>Mean unsigned error relative to CCSDT.

<sup>d</sup>Nonparallelity error relative to CCSDT.

<sup>e</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to the respective RHF- and UHF-based full CCSDT. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>f</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].

<sup>g</sup>The total CCSDT energies  $E$ , reported as  $-(199 + mE)$ , are in hartree.

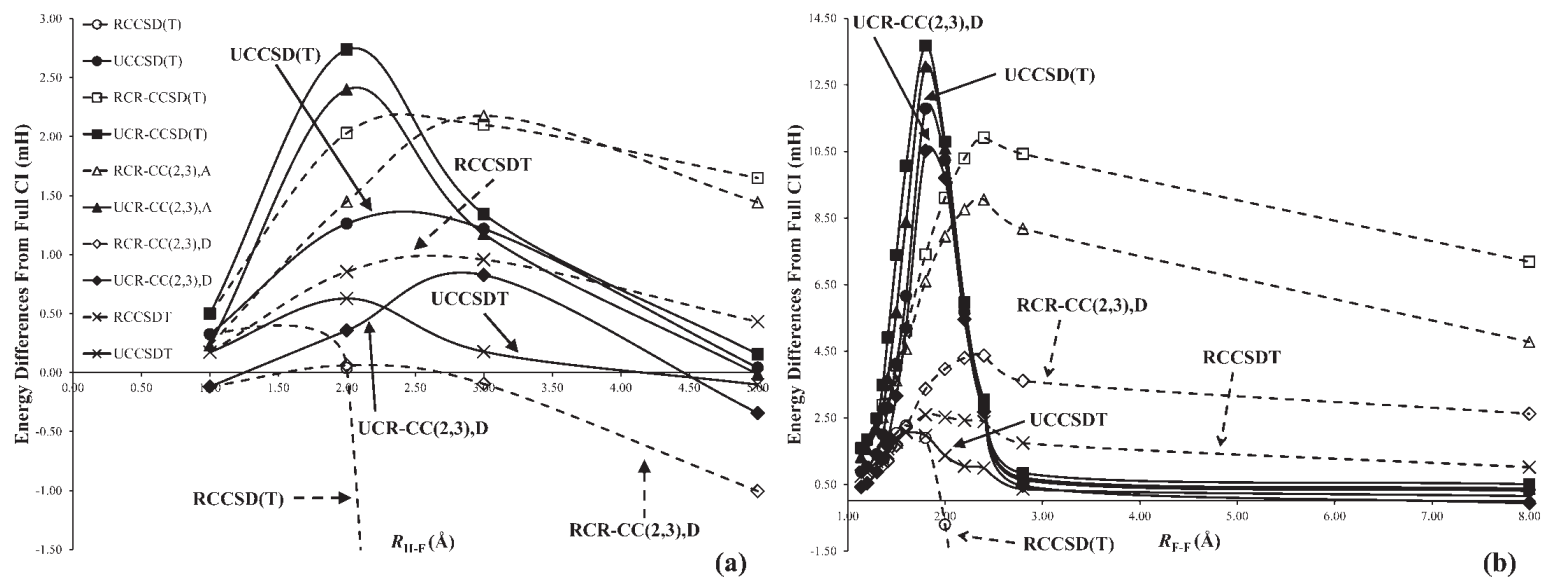


Figure 5.2: Energy differences of the various RHF- and UHF-based coupled-cluster approaches with some form of triples with respect to the full CI results for (a) the HF molecule and (b) the F<sub>2</sub> diatomic species at various bond lengths.

### 5.2.2.3 The $\text{H}_2\text{O}_2$ molecule

As mentioned above, the  $\text{H}_2\text{O}_2$  molecule serves as a representative model for O–O single bond breaking, as it is small enough that high-level CCSDT calculations can be performed for assessment of the performance of the various approximate triples CC approaches. As the CR-CC(2,3) and CCSD(T) methods are strictly size extensive these results can be used to extrapolate their performance in larger molecules involving O–O bond cleavage, where CCSDT calculations would be computationally prohibitive. We should also mention that, while the CR-CCSD(T) approach is not strictly size extensive, its violation of this property is small and thus, the conclusions drawn here, and as already shown in Ref. [425], will also be applicable when examining larger molecules. We use the cc-pVDZ basis set, freezing the core orbitals, and employing spherical  $d$  functions, keeping the rest of the geometry fixed as we stretch the O–O bond length, following the same procedure as in Ref. [425]. The results at a few O–O distances and the statistics for all of our calculations are presented in Table 5.10 and at a larger number of O–O bond lengths in the Supplementary Material and in Fig. 5.3(c).

While the RHF-based CCSD(T) results exhibit the same unphysical behavior it is well known for in bond breaking situations, it is a bit surprising to see that the UHF-based CCSD(T) results do little to improve their overall quality. Compared to their parent CCSDT calculations, the RHF-based and UHF-based CCSD(T) computations provide MSEs of -4.4 and 3.5 millihartree, respectively, and MUE values of 4.6 and 3.5 millihartree, respectively. In the previous cases, HF and  $\text{F}_2$ , the largest improvement was especially evident in the NPE results for the RHF- and UHF-based CCSD(T) calculations. For  $\text{H}_2\text{O}_2$ , however, we see there

is almost no overall improvement when the UHF reference determinant is used, resulting in NPEs of 19.6 and 17.4 millihartree, respectively. This is a good example of where even resorting to the spin contaminated UHF reference wave function will do little to improve the poor RHF-based CCSD(T) results for PESs along single bond breaking coordinates.

The CR-CCSD(T) and CR-CC(2,3),A approaches provide very similar results, with the CR-CC(2,3),A variant improving the CR-CCSD(T) calculations by about 2–3 millihartree as measured by the MSE, MUE, and NPE values when the RHF wave function is used and by about 1 millihartree for the UHF-based results, compared to their parent CCSDT energies. Comparing their RHF- and UHF-based results, we see the MSE and MUE lowering by  $\sim 1$  and increasing by 0.03 millihartree for the CR-CCSD(T) and CR-CC(2,3),A approaches, respectively. In both cases there is a significant increase in NPE when the UHF determinant is employed, by 10.6 millihartree in the CR-CCSD(T) case and 12.0 millihartree for CR-CC(2,3),A, resulting in NPE values on the same order as for the CCSD(T) calculations ( $\sim 20$  millihartree). The RHF-based CR-CC(2,3),D results, on the other hand, provide results on the order of 1 millihartree for the MSE and MUE with its NPE value remaining around 2.5 millihartree compared to its parent full CCSDT approach. Again, the effects of spin contamination for the recoupling of two doublet OH radicals into a singlet when using the UHF reference wave function is manifest in the slightly larger, though improved compared to UCCSD(T), MSE and MUE values of about 3 millihartree. The NPE for the UHF-based CR-CC(2,3),D computations is slightly less than that for UCCSD(T), but still quite large, being approximately 17 millihartree.

Table 5.10: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the H<sub>2</sub>O<sub>2</sub> molecule,<sup>a</sup> as described by the spherical cc-pVDZ basis.

| Method                   |     | 1.10  | 1.20  | 1.30  | 1.40  | $R_e^a$ | 1.50  | 1.60  | 1.70  | 1.80  | 1.90  | 2.00  | 2.10  | 2.20  | 2.30  | 2.40  | 2.50  | 2.60   | 2.70   | 2.80   | 2.90   | 3.00   | MSE <sup>b</sup> | MUE <sup>c</sup> | NPE <sup>d</sup> |
|--------------------------|-----|-------|-------|-------|-------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|--------|------------------|------------------|------------------|
| CCSD                     | RHF | 6.54  | 7.01  | 7.76  | 8.78  | 9.01    | 10.10 | 11.76 | 13.79 | 16.22 | 18.99 | 22.02 | 25.18 | 28.33 | 31.34 | 34.13 | 36.64 | 38.84  | 40.74  | 42.36  | 43.72  | 44.87  | 23.72            | 23.72            | 38.33            |
|                          | UHF | 6.54  | 7.01  | 7.76  | 8.78  | 9.01    | 10.10 | 12.91 | 16.17 | 18.86 | 29.59 | 18.80 | 16.29 | 13.31 | 10.55 | 8.30  | 6.61  | 5.45   | 4.71   | 4.29   | 4.07   | 3.96   | 10.62            | 10.62            | 25.64            |
| CC(T) <sup>e</sup>       | RHF | 0.21  | 0.21  | 0.23  | 0.26  | 0.27    | 0.30  | 0.33  | 0.34  | 0.28  | 0.08  | -0.34 | -1.07 | -2.18 | -3.68 | -5.54 | -7.69 | -10.03 | -12.45 | -14.84 | -17.13 | -19.27 | -4.37            | 4.61             | 19.60            |
|                          | UHF | 0.21  | 0.21  | 0.23  | 0.26  | 0.27    | 0.30  | 1.65  | 3.19  | 5.32  | 17.58 | 9.45  | 9.28  | 7.84  | 5.97  | 4.20  | 2.78  | 1.76   | 1.11   | 0.74   | 0.55   | 0.47   | 3.49             | 3.49             | 17.37            |
| CR(T) <sup>e</sup>       | RHF | 1.23  | 1.33  | 1.50  | 1.76  | 1.82    | 2.10  | 2.56  | 3.16  | 3.89  | 4.75  | 5.68  | 6.63  | 7.53  | 8.32  | 8.99  | 9.52  | 9.92   | 10.20  | 10.39  | 10.51  | 10.57  | 5.83             | 5.83             | 9.24             |
|                          | UHF | 1.23  | 1.33  | 1.50  | 1.76  | 1.82    | 2.10  | 4.48  | 6.96  | 9.23  | 20.67 | 11.47 | 10.56 | 8.73  | 6.65  | 4.78  | 3.30  | 2.25   | 1.59   | 1.20   | 1.01   | 0.92   | 4.93             | 4.93             | 19.75            |
| CR(2,3),A <sup>e,f</sup> | RHF | 0.78  | 0.86  | 0.99  | 1.18  | 1.22    | 1.44  | 1.79  | 2.26  | 2.84  | 3.52  | 4.26  | 4.99  | 5.66  | 6.23  | 6.67  | 6.98  | 7.18   | 7.29   | 7.33   | 7.32   | 7.28   | 4.19             | 4.19             | 6.50             |
|                          | UHF | 0.78  | 0.86  | 0.99  | 1.18  | 1.22    | 1.44  | 2.70  | 4.65  | 7.11  | 19.16 | 10.56 | 10.01 | 8.34  | 6.33  | 4.49  | 3.04  | 2.00   | 1.34   | 0.96   | 0.77   | 0.68   | 4.22             | 4.22             | 18.47            |
| CR(2,3),D <sup>e</sup>   | RHF | -0.22 | -0.24 | -0.28 | -0.28 | -0.27   | -0.18 | -0.12 | -0.03 | 0.19  | 0.48  | 0.82  | 1.18  | 1.48  | 1.74  | 1.95  | 2.09  | 2.18   | 2.23   | 2.27   | 2.28   | 2.28   | 0.93             | 1.08             | 2.52             |
|                          | UHF | -0.22 | -0.24 | -0.28 | -0.28 | -0.27   | -0.18 | 0.71  | 2.26  | 4.48  | 16.71 | 8.68  | 8.68  | 7.37  | 5.58  | 3.85  | 2.47  | 1.49   | 0.86   | 0.50   | 0.32   | 0.22   | 2.99             | 3.13             | 16.96            |

<sup>a</sup>The equilibrium geometry, with  $R_e = R_{\text{O-O}} = 1.419952\text{\AA}$  optimized at the M06-2X/MG3S level in Ref. [425]. The other geometries represent a stretching of the O–O bond without changing the other geometric parameters, and are taken from the Supporting Information to Ref. [425].

<sup>b</sup>Mean signed error relative to CCSDT.

<sup>c</sup>Mean unsigned error relative to CCSDT.

<sup>d</sup>Nonparallelity error relative to CCSDT.

<sup>e</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to the respective RHF- and UHF-based full CCSDT. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>f</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].

The differences with respect to their parent CCSDT approach are shown in Fig. 5.3(c), with all UHF-based approximate triples approaches showing even larger differences with the UCCSDT calculations, ranging from 16–20 millihartree (10–13 kcal/mol), in the intermediate O–O bond length region. Thus, it seems that any improvements gained by turning to the UHF-based CC approaches with an approximate treatment of triples, particularly in the asymptotic region of the PES along single bond breaking coordinates, are far outweighed by the inaccuracies introduced due to spin contamination at the intermediate internuclear distances. While one cannot use the RHF-based CCSD(T) approach for these problems either, the RHF-based CR-CC(2,3),D method does provide very good agreement with the full CCSDT results near equilibrium, at the intermediate bond lengths, and in the asymptotic region, where the largest errors do not exceed 2.3 millihartree (1.5 kcal/mol).

#### 5.2.2.4 The C<sub>2</sub>H<sub>6</sub> molecule

In analogy to H<sub>2</sub>O<sub>2</sub>, the C<sub>2</sub>H<sub>6</sub> molecule is a reasonable model for C–C single bond breaking, which is small enough to be able to afford the full CCSDT calculations for gauging the reliability of the various noniterative approximate triples CC methods, which can be used to extrapolate their performance in larger molecular systems where CCSDT computations are not routinely affordable. We again follow the procedure of Ref. [425], freezing the core orbitals in all CC calculations, employing the cc-pVDZ basis set and spherical *d* functions, and not allowing the geometry to relax as we stretch the internuclear C–C bond. The results at a few C–C internuclear distances and the statistics for all of our computations for C<sub>2</sub>H<sub>6</sub> are presented in Table 5.11 and Fig. 5.3(d).

The CCSD(T) results again show some improvement when the UHF determinant is used

over the RHF reference wave function, with the MSE changing from -1.0 to 2.3 millihartree in the RHF- and UHF-based CCSD(T) computations, respectively. Much like in the case of  $\text{H}_2\text{O}_2$  though, the MUE shows no change being about 2.3 millihartree in both cases. Where there is a large improvement is in their NPE values, which change from around 22 to 7 millihartree employing the RHF and UHF determinants, respectively, in the CCSD(T) calculations.

The CR-CC methods are interesting in this case in that only the CR-CCSD(T) computations show an overall improvement in the MSE and MUE values when the UHF determinant is used as a reference wave function and this improvement is only minor ( $\sim 0.3$  millihartree). In the CR-CC(2,3),A results the MSE and MUE increase by an even smaller amount (about 0.08 millihartree) and the CR-CC(2,3),D MSE and MUE increasing by 0.7 millihartree when the UHF wave function is employed. In spite of this increase when switching to the UHF reference function our CR-CC(2,3),D results still manage to improve the UCCSD(T) calculations, with our CR-CC(2,3),A and CR-CCSD(T) approaches giving comparable results to UCCSD(T). The NPEs for all UHF-based CC approaches examined in the this work, which include some form of approximate triple excitations, range from 6.6 to 8.5 millihartree, with our best UCR-CC(2,3),D approach providing the smallest NPE value, improving the UCCSD(T) approximations description of the C–C single bond breaking PES in  $\text{C}_2\text{H}_6$ . While the UCR-CC(2,3),D method improves the UCCSD(T) results, the RHF-based CR-CC(2,3),D variant provides the best overall description, with all errors not exceeding 2.2 millihartree and in some cases being significantly lower.

Table 5.11: Comparison of the energies (in millihartree) of RHF- and UHF-based CCSD and various triples-corrected CC approximations with their parent CCSDT approach for the equilibrium and displaced geometries of the  $\text{CH}_3\text{CH}_3$  molecule,<sup>a</sup> as described by the spherical cc-pVDZ basis.

| Method                   | 1.10 | 1.20  | 1.30  | 1.40  | 1.50  | $R_e^a$ | 1.60 | 1.70  | 1.80  | 1.90 | 2.00  | 2.10  | 2.20  | 2.30  | 2.40  | 2.50  | 2.60  | 2.80  | 3.00  | 3.20  | 3.50  | 4.00  | 5.00   | MSE <sup>b</sup> | MUE <sup>c</sup> | NPE <sup>d</sup> |       |
|--------------------------|------|-------|-------|-------|-------|---------|------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|------------------|------------------|------------------|-------|
| CCSD                     | RHF  | 8.54  | 8.48  | 8.52  | 8.63  | 8.81    | 8.86 | 9.06  | 9.38  | 9.78 | 10.26 | 10.85 | 11.55 | 12.38 | 13.35 | 14.47 | 15.74 | 17.16 | 20.35 | 23.82 | 27.22 | 31.68 | 36.67  | 40.58            | 15.92            | 15.92            | 32.10 |
|                          | UHF  | 8.54  | 8.48  | 8.52  | 8.63  | 8.81    | 8.86 | 9.06  | 9.38  | 9.78 | 10.26 | 10.85 | 11.56 | 13.10 | 14.75 | 16.32 | 17.55 | 18.16 | 17.07 | 13.91 | 10.71 | 7.71  | 6.41   | 6.28             | 11.07            | 11.07            | 11.88 |
| CC(T) <sup>e</sup>       | RHF  | 0.540 | 0.560 | 0.570 | 0.590 | 0.62    | 0.63 | 0.650 | 0.680 | 0.73 | 0.77  | 0.83  | 0.89  | 0.96  | 1.02  | 1.08  | 1.12  | 1.12  | 0.90  | 0.19  | -1.16 | -4.35 | -11.07 | -20.94           | -1.00            | 2.26             | 22.06 |
|                          | UHF  | 0.540 | 0.560 | 0.570 | 0.590 | 0.62    | 0.63 | 0.650 | 0.680 | 0.73 | 0.77  | 0.83  | 0.90  | 1.83  | 2.85  | 3.99  | 5.28  | 6.58  | 7.98  | 6.84  | 4.60  | 2.08  | 0.94   | 0.86             | 2.26             | 2.26             | 7.43  |
| CR(T) <sup>e</sup>       | RHF  | 1.99  | 1.99  | 2.02  | 2.07  | 2.13    | 2.15 | 2.22  | 2.33  | 2.46 | 2.62  | 2.82  | 3.07  | 3.36  | 3.71  | 4.12  | 4.59  | 5.12  | 6.29  | 7.51  | 8.59  | 9.75  | 10.46  | 10.05            | 4.41             | 4.41             | 8.47  |
|                          | UHF  | 1.99  | 1.99  | 2.02  | 2.07  | 2.13    | 2.15 | 2.22  | 2.33  | 2.46 | 2.62  | 2.82  | 3.08  | 4.68  | 6.20  | 7.59  | 8.80  | 9.70  | 9.99  | 8.16  | 5.64  | 3.00  | 1.83   | 1.73             | 4.14             | 4.14             | 8.26  |
| CR(2,3),A <sup>e,f</sup> | RHF  | 1.32  | 1.33  | 1.34  | 1.37  | 1.41    | 1.42 | 1.46  | 1.53  | 1.61 | 1.72  | 1.85  | 2.02  | 2.23  | 2.48  | 2.79  | 3.14  | 3.53  | 4.41  | 5.31  | 6.07  | 6.80  | 7.04   | 6.38             | 2.98             | 2.98             | 5.71  |
|                          | UHF  | 1.32  | 1.33  | 1.34  | 1.37  | 1.41    | 1.42 | 1.46  | 1.53  | 1.61 | 1.72  | 1.85  | 2.03  | 2.82  | 3.82  | 5.03  | 6.37  | 7.63  | 8.79  | 7.43  | 5.08  | 2.50  | 1.35   | 1.25             | 3.06             | 3.06             | 7.46  |
| CR(2,3),D <sup>e</sup>   | RHF  | 0.080 | 0.080 | 0.090 | 0.100 | 0.10    | 0.10 | 0.120 | 0.140 | 0.17 | 0.21  | 0.26  | 0.30  | 0.35  | 0.43  | 0.53  | 0.66  | 0.84  | 1.18  | 1.59  | 1.95  | 2.27  | 2.38   | 2.30             | 0.71             | 0.71             | 2.20  |
|                          | UHF  | 0.080 | 0.080 | 0.090 | 0.100 | 0.10    | 0.10 | 0.120 | 0.140 | 0.17 | 0.21  | 0.26  | 0.31  | 0.81  | 1.51  | 2.45  | 3.60  | 4.86  | 6.72  | 6.06  | 4.05  | 1.62  | 0.43   | 0.03             | 1.47             | 1.47             | 6.64  |

<sup>a</sup>The equilibrium geometry with  $R_e = R_{\text{C-C}} = 1.5227\text{\AA}$  is taken from Ref. [429]. The other geometries represent a stretching of the C–C bond without changing the other geometric parameters were taken from Ref. [425].

<sup>b</sup>Mean signed error relative to CCSDT.

<sup>c</sup>Mean unsigned error relative to CCSDT.

<sup>d</sup>Nonparallelity error relative to CCSDT.

<sup>e</sup>For the CC methods with up to triple excitations, the reported energy values, in kcal/mol, are errors relative to the respective RHF- and UHF-based full CCSDT. CC(T)  $\equiv$  CCSD(T); CR(T)  $\equiv$  CR-CCSD(T); CR(2,3),A  $\equiv$  CR-CC(2,3),A; CR(2,3),D  $\equiv$  CR-CC(2,3),D; T  $\equiv$  CCSDT.

<sup>f</sup>Equivalent to CCSD(2)<sub>T</sub> approach of Ref. [77].



Examining the energy differences of the various RHF- and UHF-based CC approaches compared to their parent CCSDT method, as shown in Fig. 5.3(d), we can again see that, while all UHF-based CC methods with an approximate form of triples provide results which approach the full CCSDT energies asymptotically, at the intermediate stretched C–C bond distances they again produce errors on the order of 7 to 10 millihartree (4–6 kcal/mol). Again, it is only the RHF-based CR-CC(2,3),D approach which is able to provide an accurate description of the PES along the C–C single bond breaking coordinate at all internuclear separations, with the largest errors not exceeding 2.4 millihartree (1.5 kcal/mol), producing results of the chemical accuracy ( $\sim 1$  kcal/mol) often touted as the standard by which to measure the performance of *ab initio* electronic structure methods.

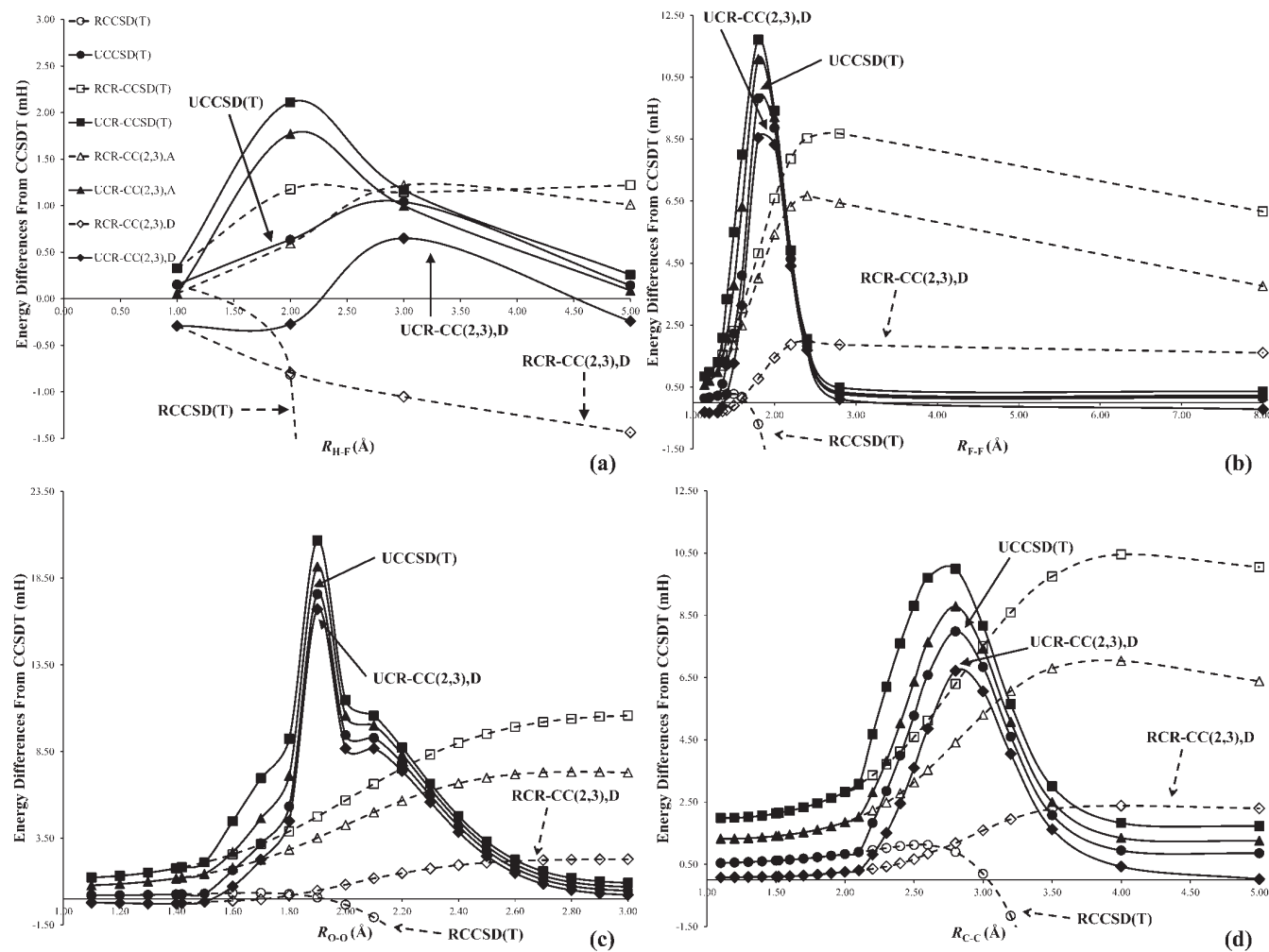


Figure 5.3: Energy differences of the various RHF- and UHF-based approximate triples coupled-cluster methods with respect to their parent CCSDT results for (a) the HF molecule, (b) the F<sub>2</sub> diatomic species, (c) the H<sub>2</sub>O<sub>2</sub> species, and (d) the C<sub>2</sub>H<sub>6</sub> polyatomic molecule at various bond lengths.

As part of the work done for this dissertation we have examined the performance of the UHF-based CR-CC approaches with approximate triples corrections, namely, CR-CCSD(T), CR-CC(2,3),A, and CR-CC(2,3),D, for describing potential energy surfaces along single bond breaking coordinates for closed-shell singlet molecules dissociating into open-shell doublet species, for the HF, F<sub>2</sub>, H<sub>2</sub>O<sub>2</sub>, and C<sub>2</sub>H<sub>6</sub> molecular species, comparing the CR-CC computations with the full CI and the parent CCSDT data as well as with their RHF-based counterparts.

Unlike the popular CCSD(T) approach, which fails to accurately describe PESs along bond breaking coordinates when the RHF determinant is used as a reference wave function and which generally is considered to be rescued by turning to the UHF reference determinant, introducing other unphysical effects, such as spin contamination and nonanalytic behavior of the PES, our CR-CC methods are not very sensitive to the reference determinant employed (RHF *vs* UHF). Our best RHF-based CR-CC(2,3),D results improve dramatically the RHF-based CCSD(T) calculations and the UHF-based CR-CC(2,3),D approach, while not as marked as in the RHF case there is still an improvement over the UHF-based CCSD(T) results for the computed PESs particularly in the region of intermediate stretched bond lengths where the spin contamination begins to render the results useless, providing errors on the order of about 10 kcal/mol, or more in some challenging cases, for all UHF-based approximate triples approaches.

From our results in this work we cannot recommend the UHF-based CC approaches with approximate triples for an accurate description of PESs along bond breaking coordinates due to the unphysical effects, which in challenging cases can be quite large, of spin contamination

providing errors on the order of some chemical reaction barriers. This effect from using UHF-based approximate triples CC methods destroys one of the important hallmark attributes that has led to their being used for challenging chemical situations, specifically that of being able to provide quantitatively correct results compared to their much more expensive parent CCSDT approaches. While one cannot use the RHF-based CCSD(T) approximation for an accurate description of PESs along bond breaking coordinates our robust RHF-based CR-CC(2,3),D variant can provide an excellent quantitative description, compared to both its parent CCSDT and the exact full CI results, of single bond breaking on a singlet PES, while avoiding the problematic behavior of the RHF- and UHF-based CCSD(T) method. In the future we would like to examine the performance of the UHF-based CR-CC approaches compared to their ROHF-based analogs for bond-breaking in open-shell species.

# Chapter 6

## Conclusions and future outlook

In this dissertation, we have described several high-level *ab initio* computational studies employing the CR-CC/CR-EOMCC and active-space CC approaches and the extensions of the EOMCC theory to open-shell systems around closed shells defining the EA-EOMCC and IP-EOMCC frameworks, demonstrating the transformative role these novel electronic structure methods, developed in our group, have played in understanding previously unexplained experiments and phenomena. Using the EA- and IP-EOMCC approaches, especially the higher-order  $3p-2h$  and  $3h-2p$  approaches invented in the Piecuch research group, we have computed the challenging electronic spectra of the CNC, C<sub>2</sub>N, N<sub>3</sub>, and NCO molecules and the photoelectron spectrum of Au<sub>3</sub><sup>-</sup>, providing for the first time an accurate interpretation of the spectrum of the latter. The CR-EOMCC formalism developed in our group played a crucial role in the discovery of the doubly excited state of azulene below the ionization threshold, which mediates the  $1 + 2'$  multiphoton ionization experiments resulting in clear Rydberg fingerprint spectra. Employing the CR-CC(2,3) methodology, developed in our group as well, we carried out a detailed investigation of the mechanism and energetics of the aerobic oxidation of methanol on Au<sub>8</sub><sup>-</sup> particle confirming the earlier DFT-based proposals that the reaction proceeds exothermically and that the rate-determining step for the reaction is the initial conversion of the methoxy species to formaldehyde. We also carried

out definitive CR-CC and active-space CC studies showing that the ground state of 1,2,3,4-cyclobutanetetraone, which is characterized by densely spaced low-lying states, is a triplet, in agreement with the recently recorded photodetachment spectrum. In addition to these applications to challenging and chemically relevant problems we discussed the development of parallel numerical energy gradients and second derivatives for fast geometry optimizations and harmonic vibrational frequency calculations at any CC/EOMCC level, allowing us to establish the geometries and relative energies of the low-energy isomers of the controversial Aug particle. We also discussed the implementation of the unrestricted Hartree-Fock-based (UHF-based) CR-CC(2,3) approach. We show that unlike the popular CCSD(T) approach, which is very sensitive to the type of the reference determinant employed in the calculations, failing in bond-breaking situations when the restricted Hartree-Fock (RHF) reference is used and displaying poor behavior at intermediate nuclear separations with UHF references, its CR-CC(2,3) counterpart provides a robust description regardless of the reference type (RHF or UHF). We also showed that the spin-adapted RHF-based CR-CC(2,3) results provide the most accurate description of the single potential energy surface for single bond breaking coordinates than all UHF-based approximate triples CC approaches in the examined cases.

In future studies, we would like to study catalytic reactions employing larger gold nanoparticles, like the non-planar Au<sub>20</sub> cluster, bimetallic catalysts, such as Pd:Au systems, as well as silver containing catalytic particles. In studying these types of systems we will likely need to make use of local correlation approaches [99–103] and/or extrapolation techniques, such as those explored as part of this dissertation. Further exploration of photochemical problems involving multiphoton processes and determining reaction pathways on the ground

and excited states, which involve locating and characterizing minimum energy crossings and seams of conical intersections, using the CIOpt package, which is able to perform these for any level of theory, should be investigated. This will present its own set of challenges and problems as the EOMCC wave function ansatz results in a non-Hermitian eigenvalue problem and which is known to have issues describing the topology and location of these minimum energy crossing points [430–433]. The development of string-based approaches to search for transition states on the ground and excited state potential energy surfaces as well as their use to locate conical intersections and to explore their seams will also be pursued in future work building on what has already been done as a part of this dissertation.

The development of MMCC corrections for the non-particle conserving EA-, IP-, DEA-, DIP-EOMCC approaches should also be pursued as it was originally proposed as part of this dissertation for the EA- and IP-EOMCC methodologies, though more than working out the relevant equations has not been done. This is also a step in the direction of extending the CC(P;Q) type approaches to the active-space EA-, IP-, DEA-, and DIP-EOMCC methods. This will hopefully allow for the extremely high-level study of challenging open-shell molecular systems, where accurate methodologies greatly help in the prediction and interpretation of their various properties.

The pursuit of higher-order parallel numerical derivative algorithms will also be undertaken in future studies along with the completion of a  $B$ -matrix routine for reducing the number of degrees of freedom to the minimum required for accurately describing molecular species. Just as this dissertation has been carried out and facilitated through the help and support of more than just the author of this dissertation so will the future studies and de-

velopments be carried out as a continued collaborative effort with the many fine scientists I have met and will yet meet and work with.



# BIBLIOGRAPHY

# BIBLIOGRAPHY

- [1] F. Coester, Nucl. Phys. **7**, 421 (1958).
- [2] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).
- [3] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
- [4] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).
- [5] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).
- [6] J. Paldus, J. Čížek, and I. Shavitt, Phys. Rev. A **5**, 50 (1972).
- [7] J. Gauss, In *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer, III, and P. R. Schreiner, vol. 1, pages 615–636 (Wiley, Chichester, 1998).
- [8] J. Paldus and X. Li, Adv. Chem. Phys. **110**, 1 (1999).
- [9] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
- [10] J. Paldus, J. Pittner, and P. Čarsky, In *Recent Progress in Coupled Cluster Methods: Theory and Applications*, edited by P. Čarsky, J. Paldus, and J. Pittner, pages 455–483 (Springer, Berlin, 2010).
- [11] D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).
- [12] G. D. Purvis, III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
- [13] J. M. Cullen and M. C. Zerner, J. Chem. Phys. **77**, 4088 (1982).
- [14] P. Piecuch and J. Paldus, Int. J. Quantum Chem. **36**, 429 (1989).

- [15] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer, III, *J. Chem. Phys.* **86**, 2881 (1987).
- [16] J. Geertsen, M. Rittby, and R. J. Bartlett, *Chem. Phys. Lett.* **164**, 57 (1989).
- [17] D. C. Comeau and R. J. Bartlett, *Chem. Phys. Lett.* **207**, 414 (1993).
- [18] J. F. Stanton and R. J. Bartlett, *J. Chem. Phys.* **98**, 7029 (1993).
- [19] H. Nakatsuji and K. Hirao, *Chem. Phys. Lett.* **47**, 569 (1977).
- [20] H. Nakatsuji, *Chem. Phys. Lett.* **59**, 362 (1978).
- [21] H. Nakatsuji, In *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczyński, vol. 2, pages 62–124 (World Scientific, Singapore, 1997).
- [22] H. Nakatsuji, *Bull. Chem. Soc. Jpn.* **78**, 1705 (2005), and references therein.
- [23] H. Monkhorst, *Int. J. Quantum Chem. Symp.* **11**, 421 (1977).
- [24] E. Dalgaard and H. Monkhorst, *Phys. Rev. A* **28**, 1217 (1983).
- [25] D. Mukherjee and P. K. Mukherjee, *Chem. Phys.* **39**, 325 (1979).
- [26] M. Takahashi and J. Paldus, *J. Chem. Phys.* **85**, 1486 (1986).
- [27] H. Koch and P. Jørgensen, *J. Chem. Phys.* **93**, 3333 (1990).
- [28] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **93**, 3345 (1990).
- [29] S. Coussan, Y. Ferro, A. Trivella, P. Roubin, R. Wieczorek, C. Manca, P. Piecuch, K. Kowalski, M. Włoch, S. A. Kucharski, and M. Musiał, *J. Phys. Chem. A* **110**, 3920 (2006).
- [30] K. Kowalski, S. Krishnamoorthy, O. Villa, J. R. Hammond, and N. Govind, *J. Chem. Phys.* **132**, 154103 (2010).

- [31] G. Fradelos, J. J. Lutz, T. A. Wesołowski, P. Piecuch, and M. Włoch, *J. Chem. Theory Comput.* **7**, 1647 (2011).
- [32] K. Kornobis, N. Kumar, P. Lodowski, M. Jaworska, P. Piecuch, J. J. Lutz, B. M. Wong, and P. M. Kozłowski, *J. Comp. Chem.* **34**, 987 (2013).
- [33] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, *J. Chem. Phys.* **128**, 134110 (2008).
- [34] S. P. A. Sauer, M. Schreiber, M. R. Silva-Junior, and W. Thiel, *J. Chem. Theory Comput.* **5**, 555 (2009).
- [35] M. R. Silva-Junior, S. P. Sauer, M. Schreiber, and W. Thiel, *Mol. Phys.* **108**, 453 (2010).
- [36] T. J. Watson, Jr., V. F. Lotrich, P. G. Szalay, A. Perera, and R. J. Bartlett, *J. Phys. Chem. A* **117**, 2569 (2013).
- [37] J. Sous, P. Goel, and M. Nooijen, *Mol. Phys.* **112**, 616 (2014).
- [38] D. Kánnár and P. G. Szalay, *J. Chem. Theory Comput.* **10**, 3757 (2014).
- [39] P. Piecuch, J. Hansen, and A. Ajala, *Mol. Phys.* (2015), in press; DOI:10.1080/00268976.2015.1076901.
- [40] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, *Int. Rev. Phys. Chem.* **21**, 527 (2002).
- [41] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P.-D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kuś, and M. Musiał, *Theor. Chem. Acc.* **112**, 349 (2004).
- [42] K. Kowalski and P. Piecuch, *J. Chem. Phys.* **115**, 643 (2001).
- [43] K. Kowalski and P. Piecuch, *Chem. Phys. Lett.* **347**, 237 (2001).
- [44] K. Kowalski and P. Piecuch, *J. Chem. Phys.* **113**, 8490 (2000).
- [45] T. Shiozaki, K. Hirao, and S. Hirata, *J. Chem. Phys.* **126**, 224106 (2007).

- [46] K. Kowalski and P. Piecuch, J. Chem. Phys. **115**, 2966 (2001).
- [47] K. Kowalski and P. Piecuch, J. Chem. Phys. **116**, 7411 (2002).
- [48] K. Kowalski and P. Piecuch, J. Chem. Phys. **120**, 1715 (2004).
- [49] M. Włoch, J. R. Gour, K. Kowalski, and P. Piecuch, J. Chem. Phys. **122**, 214107 (2005).
- [50] M. D. Lodriguito, K. Kowalski, M. Włoch, and P. Piecuch, J. Mol. Struct.: THEOCHEM **771**, 89 (2006).
- [51] M. Włoch, M. D. Lodriguito, P. Piecuch, and J. R. Gour, Mol. Phys. **104**, 2149 (2006), **104**, 2991 (2006) [Erratum].
- [52] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. **109**, 3268 (2009).
- [53] K. Kowalski, S. Hirata, M. Włoch, P. Piecuch, and T. L. Windus, J. Chem. Phys. **123**, 074319 (2005).
- [54] P. Piecuch, Mol. Phys. **108**, 2987 (2010).
- [55] J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987), **89**, 3401 (1988) [Erratum].
- [56] G. E. Scuseria and H. F. Schaefer, III, Chem. Phys. Lett. **152**, 382 (1988).
- [57] S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, J. Chem. Phys. **115**, 8263 (2001).
- [58] M. Kállay and P. Surjan, J. Chem. Phys. **113**, 1359 (2000).
- [59] S. Hirata, M. Nooijen, and R. J. Bartlett, Chem. Phys. Lett. **326**, 255 (2000).
- [60] S. Hirata, J. Chem. Phys. **121**, 51 (2004).
- [61] M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).
- [62] N. Oliphant and L. Adamowicz, J. Chem. Phys. **96**, 3739 (1992).

- [63] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. **99**, 1875 (1993).
- [64] P. Piecuch and L. Adamowicz, J. Chem. Phys. **100**, 5792 (1994).
- [65] L. Adamowicz, P. Piecuch, and K. B. Ghose, Mol. Phys. **94**, 225 (1998).
- [66] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **110**, 6103 (1999).
- [67] P. Piecuch, S. A. Kucharski, and V. Špirko, J. Chem. Phys. **111**, 6679 (1999).
- [68] J. Shen and P. Piecuch, Chem. Phys. **401**, 180 (2012).
- [69] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).
- [70] M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, J. Chem. Phys. **83**, 4041 (1985).
- [71] P. Piecuch and J. Paldus, Theor. Chim. Acta **78**, 65 (1990).
- [72] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **102**, 479 (1989).
- [73] Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **81**, 5906 (1984), **82**, 5761 (1985) [Erratum].
- [74] S. R. Gwaltney and M. Head-Gordon, Chem. Phys. Lett. **323**, 21 (2000).
- [75] S. R. Gwaltney and M. Head-Gordon, J. Chem. Phys. **115**, 2014 (2001).
- [76] S. Hirata, M. Nooijen, I. Grabowski, and R. J. Bartlett, J. Chem. Phys. **114**, 3919 (2001), **115**, 3967 (2001) [Erratum].
- [77] S. Hirata, P.-D. Fan, A. A. Auer, M. Nooijen, and P. Piecuch, J. Chem. Phys. **121**, 12197 (2004).
- [78] J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).
- [79] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 5243 (1998).

- [80] E. A. Salter and R. J. Bartlett, J. Chem. Phys. **90**, 1767 (1989).
- [81] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 1451 (1996).
- [82] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **256**, 185 (1996).
- [83] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).
- [84] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995).
- [85] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **233**, 81 (1995).
- [86] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **258**, 581 (1996).
- [87] P. Piecuch and K. Kowalski, In *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczyński, vol. 5, pages 1–104 (World Scientific, Singapore, 2000).
- [88] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
- [89] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 5644 (2000).
- [90] K. Kowalski and P. Piecuch, J. Chem. Phys. **122**, 074107 (2005).
- [91] P.-D. Fan, K. Kowalski, and P. Piecuch, Mol. Phys. **103**, 2191 (2005).
- [92] P.-D. Fan and P. Piecuch, Adv. Quantum Chem. **51**, 1 (2006).
- [93] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).
- [94] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, Chem. Phys. Lett. **418**, 467 (2006).
- [95] M. Włoch, J. R. Gour, and P. Piecuch, J. Phys. Chem. A **111**, 11359 (2007).
- [96] M. J. McGuire, K. Kowalski, and P. Piecuch, J. Chem. Phys. **117**, 3617–3624 (2002).

- [97] A. Kinal and P. Piecuch, J. Phys. Chem. A **110**, 367 (2006).
- [98] A. Kinal and P. Piecuch, J. Phys. Chem. A **111**, 734 (2007).
- [99] P. M. Kozłowski, M. Kumar, P. Piecuch, W. Li, N. P. Bauman, J. A. Hansen, P. Lodowski, and M. Jaworska, J. Chem. Theory Comput. **8**, 1870 (2012).
- [100] W. Li, J. R. Gour, P. Piecuch, and S. Li, J. Chem. Phys. **131**, 114109 (2009).
- [101] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 6721 (2010).
- [102] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
- [103] P. Arora, W. Li, P. Piecuch, J. W. Evans, M. Albao, and M. S. Gordon, J. Phys. Chem. C **114**, 12649 (2010).
- [104] P. Piecuch, S. A. Kucharski, V. Špirko, and K. Kowalski, J. Chem. Phys. **115**, 5796 (2001).
- [105] C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, J. Phys. Chem. A **110**, 1991 (2006), **111**, 4871 (2007) (Addition/Correction).
- [106] Y. Z. Song, A. Kinal, P. J. S. B. Caridade, A. J. C. Varandas, and P. Piecuch, J. Mol. Struct.: THEOCHEM **859**, 22 (2008).
- [107] X. Li and J. Paldus, J. Chem. Phys. **124**, 174101 (2006).
- [108] Y. Zhao, O. Tishchenko, J. R. Gour, W. Li, J. J. Lutz, P. Piecuch, and D. G. Truhlar, J. Phys. Chem. A **113**, 5786 (2009).
- [109] A. G. Taube, Mol. Phys. **108**, 2951 (2010).
- [110] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).
- [111] M. Z. Zgierski, S. Patchkovskii, and E. C. Lim, J. Chem. Phys. **123**, 081101 (2005).
- [112] M. Z. Zgierski, S. Patchkovskii, T. Fujiwara, and E. C. Lim, J. Phys. Chem. A **109**, 9384 (2005).



- [113] S. Nangia, D. G. Truhlar, M. J. McGuire, and P. Piecuch, *J. Phys. Chem. A* **109**, 11643 (2005).
- [114] M. Valiev and K. Kowalski, *J. Phys. Chem. A* **110**, 13106 (2006).
- [115] M. Valiev and K. Kowalski, *J. Chem. Phys.* **125**, 211101 (2006).
- [116] D. M. Chipman, *J. Chem. Phys.* **124**, 044305 (2006).
- [117] M. Z. Zgierski, S. Patchkovskii, and E. C. Lim, *Can. J. Chem.* **85**, 124 (2007).
- [118] S. Tokura, K. Yagi, T. Tsuneda, and K. Hirao, *Chem. Phys. Lett.* **436**, 30 (2007).
- [119] B. Ginovska, D. M. Camaioni, and M. Dupuis, *J. Chem. Phys.* **127**, 084309 (2007).
- [120] P.-D. Fan, M. Valiev, and K. Kowalski, *Chem. Phys. Lett.* **458**, 205 (2008).
- [121] D. M. Chipman, *J. Phys. Chem. A* **112**, 13372 (2008).
- [122] Á. Kvaran, H. Wang, K. Matthiasson, A. Bodi, and E. Jónsson, *J. Chem. Phys.* **129**, 164313 (2008).
- [123] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, *J. Chem. Phys.* **128**, 074101 (2008).
- [124] J. R. Hammond, W. A. de Jong, and K. Kowalski, *J. Chem. Phys.* **128**, 224102 (2008).
- [125] N. Govind, P. V. Sushko, W. P. Hess, M. Valiev, and K. Kowalski, *Chem. Phys. Lett.* **470**, 353 (2009).
- [126] L. A. Burns, D. Murdock, and P. H. Vaccaro, *J. Chem. Phys.* **130**, 144304 (2009).
- [127] S. I. Bokarev, E. K. Dolgov, V. A. Bataev, and I. A. Godunov, *Int. J. Quantum Chem.* **109**, 569 (2009).
- [128] K. R. Glaesemann, N. Govind, S. Krishnamoorthy, and K. Kowalski, *J. Phys. Chem. A* **114**, 8764 (2010).
- [129] J. R. Gour, P. Piecuch, and M. Włoch, *Mol. Phys.* **108**, 2633 (2010).

- [130] P. N. Day, K. A. Nguyen, and R. Pachter, *J. Chem. Theory Comput.* **6**, 2809 (2010).
- [131] K. Lopata, R. Reslan, M. Kowalska, D. Neuhauser, N. Govind, and K. Kowalski, *J. Chem. Theor. Comput.* **7**, 3686 (2011).
- [132] K. Kowalski, R. M. Olson, S. Krishnamoorthy, V. Tipparaju, and E. Aprá, *J. Chem. Theor. Comput.* **7**, 2200 (2011).
- [133] G. Fradelos, J. J. Lutz, T. A. Wesolowski, P. Piecuch, and M. Włoch, In *Advances in the Theory of Quantum Systems in Chemistry and Physics*, edited by P. E. Hoggan, E. J. Brändas, J. Maruani, P. Piecuch, and G. Delgado-Barrio, vol. 22 of *Progress in Theoretical Chemistry and Physics*, pages 219–248 (Springer, Dordrecht, 2012).
- [134] P. D. Dau, J. Su, H.-T. Liu, J.-B. Liu, D.-L. Huang, J. Li, and L.-S. Wang, *Chem. Sci.* **3**, 1137 (2012).
- [135] G. Murdachaew, M. Valiev, S. M. Kathmann, and X.-B. Wang, *J. Phys. Chem. A* **116**, 2055 (2012).
- [136] P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li, and L.-S. Wang, *J. Chem. Phys.* **137**, 064315 (2012).
- [137] N. De Silva, S. Y. Willow, and M. S. Gordon, *J. Phys. Chem. A* **117**, 11847 (2013).
- [138] J. Su, P. D. Dau, C.-F. Xu, D.-L. Huang, H.-T. Liu, F. Wei, L.-S. Wang, and J. Li, *Chem. Asian J.* **8**, 2489 (2013).
- [139] J. Su, W.-H. Xu, C.-F. Xu, W. H. E. Schwarz, and J. Li, *Inorg. Chem.* **52**, 9867 (2013).
- [140] E. Berardo, H.-S. Hu, K. Kowalski, and M. A. Zwijnenburg, *J. Chem. Phys.* **139**, 064313 (2013).
- [141] P. Tecmer, N. Govind, K. Kowalski, W. A. de Jong, and L. Visscher, *J. Chem. Phys.* **139**, 034301 (2013).
- [142] P. Piecuch, J. A. Hansen, D. Staedter, S. Faure, and V. Blanchet, *J. Chem. Phys.* **138**, 201102 (2013).

- [143] J. Su, P. D. Dau, Y.-H. Qiu, H.-T. Liu, C.-F. Xu, D.-L. Huang, L.-S. Wang, and J. Li, *Inorg. Chem.* **52**, 6617 (2013).
- [144] F. Trinter, J. B. Williams, M. Weller, M. Waitz, M. Pitzer, J. Voigtsberger, C. Schober, G. Kastirke, C. Müller, C. Goihl, P. Burzynski, F. Wiegandt, R. Wallauer, A. Kalinin, L. P. H. Schmidt, M. S. Schöffler, Y.-C. Chiang, K. Gokhberg, T. Jahnke, and R. Dörner, *Phys. Rev. Lett.* **111**, 233004 (2013).
- [145] J. J. Lutz and P. Piecuch, *Comput. Theor. Chem.* **1040-1041**, 20 (2014).
- [146] P. Zhou, J. Liu, K. Han, and G. He, *J. Comput. Chem.* **35**, 109 (2014).
- [147] G. Schoendorff and A. K. Wilson, *J. Chem. Phys.* **140**, 224314 (2014).
- [148] H.-S. Hu, K. Bhaskaran-Nair, E. Aprá, N. Govind, and K. Kowalski, *J. Phys. Chem. A* **118**, 9087 (2014).
- [149] J. Su, P. D. Dau, H.-T. Liu, D.-L. Huang, F. Wei, W. H. E. Schwarz, J. Li, and L.-S. Wang, *J. Chem. Phys.* **142**, 134308 (2015).
- [150] J. Shen and P. Piecuch, *J. Chem. Phys.* **136**, 144104 (2012).
- [151] J. Shen and P. Piecuch, *J. Chem. Theory Comput.* **8**, 4968 (2012).
- [152] C. E. Smith, R. A. King, and T. D. Crawford, *J. Chem. Phys.* **122**, 054110 (2005).
- [153] M. L. Abrams and C. D. Sherrill, *J. Chem. Phys.* **121**, 9211 (2004).
- [154] M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **102**, 3629 (1995).
- [155] M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **102**, 6735 (1995).
- [156] M. Musiał and R. J. Bartlett, *J. Chem. Phys.* **119**, 1901 (2003).
- [157] J. R. Gour, P. Piecuch, and M. Włoch, *J. Chem. Phys.* **123**, 134113 (2005).
- [158] J. R. Gour, P. Piecuch, and M. Włoch, *Int. J. Quantum Chem.* **106**, 2854 (2006).

- [159] J. R. Gour and P. Piecuch, *J. Chem. Phys.* **125**, 234107 (2006).
- [160] Y. Ohtsuka, P. Piecuch, J. R. Gour, M. Ehara, and H. Nakatsuji, *J. Chem. Phys.* **126**, 164111 (2007).
- [161] M. Kamiya and S. Hirata, *J. Chem. Phys.* **125**, 074111 (2006).
- [162] P.-D. Fan, M. Kamiya, and S. Hirata, *J. Chem. Theor. Comput.* **3**, 1036 (2007).
- [163] R. J. Bartlett and J. F. Stanton, In *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd, vol. 5, pages 65–169 (VCH Publishers, New York, 1994).
- [164] M. Nooijen and J. G. Snijders, *Int. J. Quantum Chem. Symp.* **26**, 55 (1992).
- [165] M. Nooijen and J. G. Snijders, *Int. J. Quantum Chem.* **48**, 15 (1993).
- [166] J. F. Stanton and J. Gauss, *J. Chem. Phys.* **101**, 8938 (1994).
- [167] M. Musiał, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **118**, 1128 (2003).
- [168] M. Musiał and R. J. Bartlett, *Chem. Phys. Lett.* **384**, 210 (2004).
- [169] Y. J. Bomble, J. C. Saeh, J. F. Stanton, P. G. Szalay, and M. Kállay, *J. Chem. Phys.* **122**, 154107 (2005).
- [170] M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **106**, 6441 (1997).
- [171] M. Wladyslawski and M. Nooijen, In *Low-Lying Potential Energy Surfaces*, edited by M. R. Hoffmann and K. G. Dyall, vol. 828 of *ACS Symposium Series*, pages 65–92 (American Chemical Society, Washington, D.C., 2002).
- [172] M. Nooijen, *Int. J. Mol. Sci.* **3**, 656 (2002).
- [173] R. J. Bartlett, *Mol. Phys.* **108**, 2905 (2010).
- [174] M. Musiał, A. Perera, and R. J. Bartlett, *J. Chem. Phys.* **134**, 114108 (2011).

- [175] M. Musiał, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Theory Comput.* **7**, 3088 (2011).
- [176] K. W. Sattelmeyer, H. F. Schaefer, III, and J. F. Stanton, *Chem. Phys. Lett.* **378**, 42 (2003).
- [177] J. Shen and P. Piecuch, *J. Chem. Phys.* **138**, 194102 (2013).
- [178] J. Shen and P. Piecuch, *Mol. Phys.* **112**, 868 (2014).
- [179] M. Ehara, J. R. Gour, and P. Piecuch, *Mol. Phys.* **107**, 871 (2009).
- [180] J. Hansen, P. Piecuch, J. Lutz, and J. Gour, *Phys. Scr.* **84**, 028110 (2011).
- [181] J. A. Hansen, M. Ehara, and P. Piecuch, *J. Phys. Chem. A* **117**, 10416 (2013).
- [182] J. Hansen, N. Bauman, J. Shen, W. Borden, and P. Piecuch In preparation.
- [183] J. A. Hansen, P. Piecuch, and B. G. Levine, *J. Chem. Phys.* **139**, 091101 (2013).
- [184] N. P. Bauman, J. A. Hansen, M. Ehara, and P. Piecuch, *J. Chem. Phys.* **141**, 101102 (2014).
- [185] T. H. Dunning, Jr., *J. Chem. Phys.* **53**, 2823 (1970).
- [186] T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- [187] T. Nakajima and H. Nakatsuji, *Chem. Phys. Lett.* **79**, 280 (1997).
- [188] M. Ishida, K. Toyota, M. Ehara, and H. Nakatsuji, *Chem. Phys. Lett.* **347**, 493 (2001).
- [189] M. Ishida, K. Toyota, M. Ehara, H. Nakatsuji, and M. J. Frisch, *J. Chem. Phys.* **120**, 2593 (2004).
- [190] 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).

- [191] M. S. Gordon and M. W. Schmidt, In *Theory and Applications of Computational Chemistry: The First Forty Years*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria, pages 1167–1190 (Elsevier, Amsterdam, 2005).
- [192] J. R. Gour, P. Piecuch, and M. Włoch, *Mol. Phys.* **108**, 2633 (2010).
- [193] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- [194] D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995).
- [195] G. Herzberg. *Molecular Spectra and Molecular Structure III: Electronic Spectra and Electronic Structure of Polyatomic Molecules* (Van Nostrand, London, 1967).
- [196] A. Merer and D. Travis, *Can. J. Phys.* **44**, 353 (1966).
- [197] A. Merer and D. Travis, *Can. J. Phys.* **43**, 1795 (1965).
- [198] R. Pd. and P. Chandra, *J. Chem. Phys.* **114**, 1589 (2001).
- [199] G. Hutchings, *J. Catal.* **96**, 292 (1985).
- [200] M. Haruta, T. Kobayashi, H. Sano, and N. Yamada, *Chem. Lett.* **16**, 405 (1987).
- [201] A. S. K. Hashmi and G. J. Hutchings, *Angewandte Chemie International Edition* **45**, 7896 (2006).
- [202] T. Tsukuda, H. Tsunoyama, and H. Sakurai, *Chem. Asian J.* **6**, 736 (2011).
- [203] T. Takei, T. Akita, I. Nakamura, T. Fujitani, M. Okumura, K. Okazaki, J. Huang, T. Ishida, and M. Haruta, In *Advances in Catalysis*, edited by B. C. Gates and F. C. Jentoft, vol. 55, pages 1 – 126 (Academic Press, San Diego, 2012).
- [204] P. Pyykkö, *Angew. Chem., Int. Ed.* **43**, 4412 (2004).
- [205] P. Pyykkö, *Inorg. Chim. Acta* **458**, 4113 (2005).
- [206] H. Häkkinen, *Chem. Soc. Rev.* **37**, 1847 (2008).

- [207] P. Pyykkö, Chem. Soc. Rev. **37**, 1967 (2008).
- [208] R. Wesendrup, T. Hunt, and P. Schwerdtfeger, J. Chem. Phys. **112**, 9356 (2000).
- [209] M. P. Johansson, A. Lechtken, D. Schooss, M. M. Kappes, and F. Furche, Phys. Rev. A **77**, 053202 (2008).
- [210] A. A. Rusakov, E. Rykova, G. E. Scuseria, and A. Zaitsevskii, J. Chem. Phys. **127**, 164322 (2007).
- [211] S. A. Varganov, R. M. Olson, M. S. Gordon, and H. Metiu, J. Chem. Phys. **119**, 2531 (2003).
- [212] Y. Shen and J. J. BelBruno, J. Phys. Chem. A **109**, 512 (2005).
- [213] S. A. Varganov, R. M. Olson, M. S. Gordon, and H. Metiu, J. Chem. Phys. **120**, 5169 (2004).
- [214] R. M. Olson, S. Varganov, M. S. Gordon, H. Metiu, S. Chretien, P. Piecuch, K. Kowalski, S. A. Kucharski, and M. Musial, J. Am. Chem. Soc. **127**, 1049 (2005).
- [215] M. Diefenbach and K. S. Kim, J. Phys. Chem. B **110**, 21639 (2006).
- [216] Y. K. Han, J. Chem. Phys. **124**, 024316 (2006).
- [217] R. M. Olson and M. Gordon, J. Chem. Phys. **126**, 214310 (2007).
- [218] D. A. Götz, R. Schäfer, and P. Schwerdtfeger, J. Comp. Chem. **34**, 1975 (2013).
- [219] J. V. Koppen, M. Hapka, M. M. Szcześniak, and G. Chałasiński, J. Chem. Phys. **137**, 114302 (2012).
- [220] S. Lecoultrre, A. Rydlo, C. Félix, J. Buttet, S. Gilb, and W. Harbich, J. Chem. Phys. **134**, 074302 (2011).
- [221] K. J. Taylor, C. L. Pettiette-Hall, O. Cheshnovsky, and R. E. Smalley, J. Chem. Phys. **96**, 3319 (1992).

- [222] H. Handschuh, G. Ganteför, P. S. Bechthold, and W. Eberhardt, *J. Chem. Phys.* **100**, 7093 (1994).
- [223] H. Häkkinen, B. Yoon, U. Landman, X. Li, H.-J. Zhai, and L.-S. Wang, *J. Phys. Chem. A* **107**, 6168 (2003).
- [224] C. W. Bauschlicher, Jr., *Chem. Phys. Lett.* **156**, 91 (1989).
- [225] C. W. Bauschlicher, Jr., S. R. Langhoff, and H. Partridge, *J. Chem. Phys.* **91**, 2412 (1989).
- [226] K. K. Das and K. Balasubramanian, *Chem. Phys. Lett.* **176**, 571 (1991).
- [227] K. Balasubramanian and K. K. Das, *Chem. Phys. Lett.* **186**, 577 (1991).
- [228] A. A. Rusakov and A. Zaitsevskii, *Cent. Eur. J. Phys.* **6**, 771 (2008).
- [229] K. A. Peterson and C. Puzzarini, *Theor. Chem. Acc.* **114**, 283 (2005).
- [230] D. Figgen, G. Rauhut, M. Dolg, and H. Stoll, *Chem. Phys.* **311**, 227 (2005).
- [231] M. Dolg, K. A. Peterson, P. Schwerdtfeger, and H. Stoll (2009). [Http://www.tc.uni-koeln.de/PP/index.en.html](http://www.tc.uni-koeln.de/PP/index.en.html).
- [232] T. Andersen, H. K. Haugen, and H. Hotop, *J. Phys. Chem. Ref. Data* **28**, 1511 (1999).
- [233] NIST (2013). [Http://www.nist.gov/pml/data/asd.cfm](http://www.nist.gov/pml/data/asd.cfm).
- [234] K. Dyal, I. Grant, C. Johnson, F. Parpia, and E. Plummer, *Comput. Phys. Commun.* **55**, 425 (1989).
- [235] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, *Mol. Phys.* **104**, 1039 (2006).
- [236] G. A. Bishea and M. D. Morse, *J. Chem. Phys.* **95**, 8779 (1991).
- [237] H. Partridge, C. W. B. Jr., and S. R. Langhoff, *Chem. Phys. Lett.* **175**, 531 (1990).



- [238] P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musiał, *Comp. Phys. Commun.* **149**, 71 (2002).
- [239] R. Haufler, L.-S. Wang, L. Chibante, C. Jin, J. Conceicao, Y. Chai, and R. Smalley, *Chem. Phys. Lett.* **179**, 449 (1991).
- [240] A. Amrein, R. Simpson, and P. Hackett, *J. Chem. Phys.* **94**, 4663 (1991).
- [241] J. A. Howard, R. Sutcliffe, and B. Mile, *J. Chem. Soc., Chem. Commun.* **23**, 1449 (1983).
- [242] L. Meissner and R. Bartlett, *jcp* **102**, 7490 (1995).
- [243] V. Blanchet, K. Raffael, G. Turri, B. Chatel, B. Girard, I. Garcia, I. Wilkinson, and B. Whitaker, *J. Chem. Phys.* **128**, 164318 (2008).
- [244] M. Kasha, *Discuss. Faraday Soc.* **9**, 14 (1950).
- [245] E. W.-G. Diau, S. De Feyter, and A. H. Zewail, *J. Chem. Phys.* **110**, 9785 (1999).
- [246] O. K. Abou-Zied, D. R. M. Demmer, S. C. Wallace, and R. P. Steer, *Chem. Phys. Lett.* **266**, 75 (1997).
- [247] Y. Hirata and E. C. Lim, *J. Chem. Phys.* **69**, 3292 (1978).
- [248] T. Makinoshima, M. Fujitsuka, M. Sasaki, Y. Araki, O. Ito, S. Ito, and N. Morita, *J. Phys. Chem. A* **108**, 368 (2004).
- [249] P. M. Weber and N. Thantu, *Chem. Phys. Lett.* **197**, 556 (1992).
- [250] M. Fujii, T. Ebata, N. Mikami, and M. Ito, *Chem. Phys.* **77**, 191 (1983).
- [251] P. Foggi and F. V. R. Neuwahl, *J. Phys. Chem. A* **107**, 1689 (2003).
- [252] J. R. Cable and A. C. Albrecht, *J. Chem. Phys.* **84**, 1969 (1986).
- [253] L. Ciano, P. Foggi, and P. Remigio-Salvi, *J. Photochem. Photobiol. B* **105**, 129 (1997).

- [254] M. Damm, F. Deckert, H. Hippler, and J. Troe, *J. Phys. Chem.* **95**, 2005 (1991).
- [255] H. Hippler, L. Lindemann, and J. Troe, *J. Chem. Phys.* **83**, 3906 (1985).
- [256] U. Hold, T. Lenzer, K. Luther, and A. C. Symonds, *J. Chem. Phys.* **119**, 11192 (2003).
- [257] K. Raffael, V. Blanchet, B. Chatel, G. Turri, B. Girard, I. A. Garcia, I. Wilkinson, and B. J. Whitaker, *Chem. Phys. Lett.* **460**, 59 (2008).
- [258] M. Deleuze, *J. Chem. Phys.* **116**, 7012 (2002).
- [259] N. Kuthirummal and P. M. Weber, *J. Mol. Struct.* **787**, 163 (2006).
- [260] W. Gerhartz and J. Michl, *J. Am. Chem. Soc.* **100**, 6877 (1978).
- [261] N. Verdal, S. A. Rivera, and B. S. Hudson, *Chem. Phys. Lett.* **437**, 38 (2007).
- [262] A. Murakami, T. Kobayashi, A. Goldberg, and S. Nakamura, *J. Chem. Phys.* **120**, 1245 (2004).
- [263] P. Kozlowski, G. Rauhut, and P. Pulay, *J. Chem. Phys.* **103**, 5650 (1995).
- [264] J. M. L. Martin, J. El-Yazal, and J.-P. François, *J. Phys. Chem.* **100**, 15358 (1996).
- [265] A. Hanson, *Acta Crystallogr.* **19**, 19 (1965).
- [266] O. Bastiansen and J. L. Derissen, *Acta Chem. Scand.* **20**, 1319 (1966).
- [267] A. D. Hunter, E. E. Burnell, and T. C. Wong, *J. Mol. Struct.* **99**, 159 (1983).
- [268] S. Huber, G. Grassi, and A. Bauder, *Mol. Phys.* **103**, 1395 (2005).
- [269] K. Ruedenberg, M. Schmidt, M. Gilbert, and S. Eliot, *Chem. Phys.* **71**, 41 (1982).
- [270] B. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).

- [271] K. Anderson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Woliński, *J. Phys. Chem.* **94**, 5483 (1990).
- [272] K. Anderson, P.-Å. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96**, 1218 (1992).
- [273] R. Pariser, *J. Chem. Phys.* **25**, 1112 (1956).
- [274] N. Kuthirummal and P. M. Weber, *Chem. Phys. Lett.* **378**, 647 (2003).
- [275] N. Thire, D. Staedter, R. Cireasa, V. Blanchet, and S. T. Pratt, *Phys. Chem. Chem. Phys.* **13**, 18485 (2011).
- [276] S. Deb and P. M. Weber, *Annu. Rev. Phys. Chem.* **62**, 19 (2011).
- [277] J. O. Johansson, G. G. Henderson, F. Remacle, and E. E. B. Campbell, *Phys. Rev. Lett.* **108**, 173401 (2012).
- [278] P. Hemberger, J. Kohler, I. Fischer, G. Piani, L. Poisson, and J.-M. Mestdagh, *Phys. Chem. Chem. Phys.* **14**, 6173 (2012).
- [279] M. P. Minitti, J. D. Cardoza, and P. M. Weber, *J. Phys. Chem. A* **110**, 10212 (2006).
- [280] M. P. Minitti and P. M. Weber, *Phys. Rev. Lett.* **98**, 253004 (2007).
- [281] C. Christensen and J. Nørskov, *Science* **327**, 278 (2010).
- [282] A. Corma, A. Leyva-Pérez, and M. J. Sabater, *Chem. Rev.* **111**, 1657 (2011).
- [283] Z. Li, C. Brouwer, and C. He, *Chem. Rev.* **108**, 3239 (2008).
- [284] R. Sardar, A. M. Funston, P. Mulvaney, and R. W. Murray, *Langmuir* **25**, 13840 (2009).
- [285] B. Hammer and J. Nørskov, In *Impact of Surface Science on Catalysis*, edited by H. K. Bruce C. Gates, vol. 45, pages 71 – 129 (Academic Press, 2000).
- [286] T. Ishida, M. Nagaoka, T. Akita, and M. Haruta, *Chem. Eur. J.* **14**, 8456 (2008).

- [287] M. Okumura, Y. Kitagawa, T. Kawakami, and M. Haruta, *Chem. Phys. Lett.* **459**, 133 (2008).
- [288] M. Conte, H. Miyamura, S. Kobayashi, and V. Chechik, *J. Amer. Chem. Soc.* **131**, 7189 (2009).
- [289] M. Conte, H. Miyamura, S. Kobayashi, and V. Chechik, *Chem. Commun.* **46**, 145 (2010).
- [290] Y. B. Ge, M. S. Gordon, and P. Piecuch, *J. Chem. Phys.* **127**, 174106 (2007).
- [291] Y. B. Ge, M. S. Gordon, P. Piecuch, M. Włoch, and J. R. Gour, *J. Phys. Chem. A* **112**, 11873 (2008).
- [292] G. R. Magoon, J. Aguilera-Iparraguirre, W. H. Green, J. J. Lutz, P. Piecuch, H.-W. Wong, and O. O. Oluwole, *Int. J. Chem. Kin.* **44**, 179 (2012).
- [293] R. Coquet, K. L. Howard, and D. J. Willock, *Chem. Soc. Rev.* **37**, 2046 (2008).
- [294] L. Molina and B. Hammer, *Appl. Catal. A* **291**, 21 (2005).
- [295] H. Nakatsuji, Z.-M. Hu, H. Nakai, and K. Ikeda, *Surf. Sci.* **387**, 328 (1997).
- [296] M. Okumura, Y. Kitagawa, and T. Kawakami, *Int. J. Quantum Chem.* **110**, 2903 (2010).
- [297] A. Lyalin and T. Taketsugu, *J. Phys. Chem. C* **113**, 12930 (2009).
- [298] A. Lyalin and T. Taketsugu, *J. Phys. Chem. Lett.* **1**, 1752 (2010).
- [299] K. Bobuatong, S. Karanjit, R. Fukuda, M. Ehara, and H. Sakurai, *Phys. Chem. Chem. Phys.* **14**, 3103 (2012).
- [300] S. Karanjit, K. Bobuatong, R. Fukuda, M. Ehara, and H. Sakurai, *Int. J. Quantum Chem.* **113**, 428 (2013).
- [301] Y. Zhao and D. Truhlar, *J. Chem. Phys.* **125**, 194101 (2006).

- [302] B. Salisbury, W. Wallace, and R. Whetten, Chem. Phys. **262**, 131 (2000).
- [303] H. Zhang, T. Watanabe, M. Okumura, M. Haruta, and N. Toshima, Nat. Mater. **11**, 49 (2012).
- [304] H. Zhang, M. Okumura, and N. Toshima, J. Phys. Chem. C **115**, 14883 (2011).
- [305] R. N. Dhital, C. Kamonsatikul, E. Somsook, K. Bobuatong, M. Ehara, S. Karanjit, and H. Sakurai, J. Am. Chem. Soc. **134**, 20250 (2012).
- [306] A. D. Becke, J. Chem. Phys. **84**, 4524 (1986).
- [307] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
- [308] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
- [309] S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
- [310] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [311] J.-D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**, 6615 (2008).
- [312] H. Tsunoyama, N. Ichikuni, H. Sakurai, and T. Tsukuda, J. Am. Chem. Soc. **131**, 7086 (2009).
- [313] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
- [314] M. Guest and V. Saunders, Mol. Phys. **28**, 819 (1974).
- [315] P. Hay and W. Wadt, J. Chem. Phys. **82**, 270 (1985).
- [316] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
- [317] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
- [318] T. Clark, J. Chandrasekhar, G. Spitznagel, and P. Schleyer, J. Comput. Chem. **4**, 294 (1983).

- [319] L. Roy, J. Hay, and R. Martin, *J. Chem. Theory Comput.* **4**, 1029 (2008).
- [320] K. Raghavachari, J. Binkley, R. Seeger, and J. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- [321] A. Ehlers, M. Böhme, S. Dapprich, A. Gobbi, A. Höllwarth, V. Jonas, K. Köller, R. Stegmann, A. Veldkamp, and G. Frenking, *Chem. Phys. Lett.* **208**, 111 (1993).
- [322] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, *Comp. Phys. Commun.* **181**, 1477 (2010).
- [323] J. J. Zheng, J. R. Gour, J. J. Lutz, M. Włoch, P. Piecuch, and D. G. Truhlar, *J. Chem. Phys.* **128**, 044108 (2008).
- [324] J. Lutz and P. Piecuch, *J. Chem. Phys.* **128**, 154116 (2008).
- [325] R. Gleiter, I. Hyla-Kryspin, and K.-H. Pfeifer, *J. Org. Chem.* **60**, 5878 (1995).
- [326] H. Jiao, G. Frapper, J.-F. Halet, and J.-Y. Saillard, *J. Phys. Chem. A* **105**, 5945 (2001).
- [327] R. G. X. Zhou, D.A. Hrovat and W. Borden, *Mol. Phys.* **107**, 863 (2009).
- [328] R. G. X. Zhou, D.A. Hrovat and W. Borden, *J. Phys. Chem. A* **114**, 1304 (2010).
- [329] J.-C. Guo, G.-L. Hou, S.-D. Li, and X.-B. Wang, *J. Phys. Chem. Lett.* **3**, 304 (2012).
- [330] X. Bao, D. A. Hrovat, W. T. Borden, and X.-B. Wang, *J. Am. Chem. Soc.* **135**, 4291 (2013).
- [331] A. Perera, R. W. Molt, Jr., V. F. Lotrich, and R. Bartlett, *Theor. Chem. Acc.* **133**, 1514 (2014).
- [332] J. M. Foster and S. F. Boys, *Rev. Mod. Phys.* **32**, 305 (1960).
- [333] J. F. Harrison and L. C. Allen, *J. Am. Chem. Soc.* **91**, 807 (1969).
- [334] C. F. Bender and H. F. Schaefer III, *J. Am. Chem. Soc.* **92**, 4984 (1970).

- [335] G. Herzberg, Proc. R. Soc. London, Ser. A **262**, 291 (1961).
- [336] S. J. Cole, G. D. P. III, and R. J. Bartlett, Chem. Phys. Lett. **113**, 271 (1985).
- [337] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, J. Chem. Phys. **87**, 387 (1987).
- [338] E. A. Carter and W. A. Goddard, J. Chem. Phys. **86**, 862 (1987).
- [339] A. D. McLean, P. R. Buenker, R. M. Escibano, and P. Jensen, J. Chem. Phys. **87**, 2166 (1987).
- [340] D. C. Comeau, I. Shavitt, P. Jensen, and P. R. Bunker, J. Chem. Phys. **90**, 6491 (1989).
- [341] D. B. Knowles, J. Ramon, A. Collado, G. Hirsch, and R. J. Buekner, J. Chem. Phys. **92**, 585 (1990).
- [342] J. D. Watts and R. J. Bartlett, J. Chem. Phys. **93**, 6104 (1990).
- [343] K. Hirao, Chem. Phys. Lett. **196**, 397 (1992).
- [344] M. Hoffmann, Chem. Phys. Lett. **195**, 127 (1992).
- [345] P. J. Knowles, C. Hampel, and H.-J. Werner, J. Chem. Phys. **99**, 5219 (1993).
- [346] C. D. Sherrill, T. J. V. Huis, T. J. Yamaguchi, and H. F. Schaefer, III, J. Mol. Struct.: THEOCHEM **400**, 139 (1997).
- [347] C. D. Sherrill, M. L. Leininger, T. J. V. Huis, and H. F. Schaefer, III, J. Chem. Phys. **108**, 1040 (1998).
- [348] X. Li and J. Paldus, Collect. Czech. Chem. Commun. **63**, 1381 (1998).
- [349] X. Li and J. Paldus, J. Chem. Phys. **129**, 174101 (2008).
- [350] Y. G. Khait, W. Jiang, and M. Hoffmann, Int. J. Quantum Chem. **109**, 1855 (2009).

- [351] P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, and C. J. Umrigar, J. Chem. Phys. **131**, 124103 (2009).
- [352] X. Li, P. Piecuch, and J. Paldus, Chem. Phys. Lett. **224**, 267 (1994).
- [353] P. Piecuch, X. Li, and J. Paldus, Chem. Phys. Lett. **230**, 377 (1994).
- [354] H. F. Schaefer III, Science **231**, 1100 (1986).
- [355] J. F. Harrison, In *Advances in the Theory of Atomic and Molecular Systems: Conceptual and Computational Advances in Quantum Chemistry*, edited by P. Piecuch, J. Maruani, G. Delgado-Barrio, and S. Wilson, vol. 19 of *Progress in Theoretical Chemistry and Physics*, pages 33–43 (Springer, Dordrecht, 2009).
- [356] X. Zhou, D. Hrovat, and W. Borden, Org. Lett. **10**, 893 (2008).
- [357] X. Bao, X. Zhou, C. Flener Lovitt, A. Venkatraman, D. A. Hrovat, R. Gleiter, R. Hoffmann, and W. T. Borden, J. Am. Chem. Soc. **134**, 10259 (2012).
- [358] W. T. Borden, Isr. J. Chem. (2015), in press; Publication Date (Web) : February 20, 2015; DOI:10.1002/ijch.201400170.
- [359] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).
- [360] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).
- [361] K. B. Ghose, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **103**, 9331 (1995).
- [362] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6769 (1992).
- [363] B. G. Levine, J. D. Coe, and T. J. Martínez, J. Phys. Chem. **112B**, 405 (2008).
- [364] Energy equivalents corresponding to the unique harmonic vibrational frequencies, scaled by 0.97, of the four electronic states of C<sub>4</sub>O<sub>4</sub>, computed at the UB3LYP/aug-cc-pVTZ level (in kJ/mol) are as follows.  $8\pi(^1A_{1g})$ : (B<sub>2u</sub>) 1.05, (A<sub>2u</sub>) 2.71, (B<sub>2g</sub>) 3.13, (E<sub>u</sub>) 3.41, (E<sub>g</sub>) 6.25, (B<sub>1g</sub>) 6.82, (A<sub>1g</sub>) 7.20, (B<sub>2u</sub>) 8.53, (E<sub>u</sub>) 8.64, (A<sub>2g</sub>) 8.75, (B<sub>2g</sub>) 11.10, (E<sub>u</sub>) 20.04, (A<sub>1g</sub>) 21.87, (B<sub>1g</sub>) 22.17.  $9\pi(^3B_{2u})$ : (B<sub>2u</sub>) 1.05, (A<sub>2u</sub>) 2.72, (B<sub>2g</sub>) 3.15, (E<sub>u</sub>) 3.42, (E<sub>g</sub>) 6.17, (B<sub>1g</sub>) 6.82, (A<sub>1g</sub>) 7.23, (B<sub>2u</sub>) 8.58, (A<sub>2g</sub>) 8.76, (E<sub>u</sub>)



8.76, ( $B_{2g}$ ) 11.10, ( $E_u$ ) 19.96, ( $A_{1g}$ ) 21.79, ( $B_{1g}$ ) 22.09.  $9\pi(^1B_{2u})$ : ( $B_{2u}$ ) 0.92, ( $A_{2u}$ ) 2.44, ( $B_{2g}$ ) 2.83, ( $E_u$ ) 3.25, ( $E_g$ ) 4.54, ( $A_{1g}$ ) 6.99, ( $B_{1g}$ ) 7.17, ( $A_{2g}$ ) 8.93, ( $B_{2g}$ ) 9.68, ( $B_{2u}$ ) 10.41, ( $E_u$ ) 11.40, ( $E_u$ ) 21.06, ( $A_{1g}$ ) 21.71, ( $B_{1g}$ ) 22.28.  $10\pi(^1A_{1g})$ : ( $B_{2u}$ ) 1.03, ( $A_{2u}$ ) 2.94, ( $B_{2g}$ ) 3.38, ( $E_u$ ) 3.44, ( $B_{1g}$ ) 6.42, ( $B_{2u}$ ) 6.43, ( $E_u$ ) 6.69, ( $A_{1g}$ ) 7.35, ( $E_g$ ) 7.58, ( $A_{2g}$ ) 8.50, ( $B_{2g}$ ) 12.23, ( $E_u$ ) 20.28, ( $A_{1g}$ ) 22.23, ( $B_{1g}$ ) 22.33.

- [365] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, and T. Thorsteinsson. MOLPRO, version 2006.1, a package of ab initio programs.
- [366] L. Adamowicz, W. Laidig, and R. Bartlett, *Int. J. Quantum Chem. Symp.* **18**, 245 (1984).
- [367] E. A. Salter, G. W. Trucks, and R. J. Bartlett, *J. Chem. Phys.* **90**, 1752 (1989).
- [368] J. Gauss, J. Stanton, and R. Bartlett, *J. Chem. Phys.* **95**, 2623 (1991).
- [369] J. D. Watts, J. Gauss, and R. J. Bartlett, *Chem. Phys. Lett.* **200**, 1 (1992).
- [370] J. D. Watts, J. Gauss, and R. J. Bartlett, *J. Chem. Phys.* **98**, 8718 (1993).
- [371] J. F. Stanton, *J. Chem. Phys.* **99**, 8840 (1993).
- [372] A. G. Taube and R. J. Bartlett, *J. Chem. Phys.* **128**, 044110 (2008).
- [373] A. G. Taube and R. J. Bartlett, *J. Chem. Phys.* **128**, 044111 (2008).
- [374] J. F. Stanton and J. Gauss, *Int. Rev. Phys. Chem.* **19**, 61 (2000).
- [375] M. Harding, T. Metzroth, and J. Gauss, *J. Chem. Theory Comput.* **4**, 64 (2008).
- [376] T. J. Lee and C. E. Dateo, *J. Chem. Phys.* **107**, 10373 (1997).
- [377] J. M. L. Martin, T. J. Lee, and P. R. Taylor, *J. Chem. Phys.* **108**, 676 (1998).

- [378] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez, *J. Phys. Chem. B* **112**, 12559 (2008).
- [379] S. K. Burger and P. W. Ayers, *J. Chem. Phys.* **133**, 034116 (2010).
- [380] P. Pulay, *Mol. Phys.* **17**, 197 (1969).
- [381] T. H. Dunning, Jr. and P. J. Hay, In *Practical Methods of Optimization* (Wiley; 2 edition, New Jersey, 2000).
- [382] M. Hestenes and E. Stiefel, *J. Research NIST* **49**, 409 (1952).
- [383] J. A. Nelder and R. Mead, *Computer J.* **7**, 308 (1965).
- [384] C. G. Broyden, *J. Inst. Math. Appl.* **6**, 79 (1970).
- [385] C. G. Broyden, *J. Inst. Math. Appl.* **6**, 222 (1970).
- [386] R. Fletcher, *Computer J.* **13**, 317 (1970).
- [387] D. Goldfarb, *Math. Comp.* **24**, 23 (1970).
- [388] D. F. Shanno, *Math. Comp.* **24**, 647 (1970).
- [389] A. R. Conn, N. I. M. Gould, and P. L. Toint, *SIAM J. Numer. Anal.* **25**, 433 (1988), **26**, 764 (1989) [Erratum].
- [390] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Math. Comp.* **50**, 399 (1988).
- [391] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Math. Prog.* **50**, 177 (1991).
- [392] H. F. Khalfan, R. H. Byrd, and R. B. Schnabel, *SIAM J. Optim.* **3**, 1 (1993).
- [393] A. Mitin, *J. Comput. Chem.* **19**, 1877 (1998).
- [394] J. Dennis and H. Wolkowicz, *SIAM J. Numer. Anal.* **30**, 1291 (1993).

- [395] H. Wolkowicz, Math. Oper. Res. **19**, 815 (1994).
- [396] W. J. Leong and M. A. Hassan, Appl. Math. Comput. **218**, 413 (2011).
- [397] D. Liu and J. Nocedal, Math. Program. B **45**, 503 (1989).
- [398] A. Schäfer, H. Horn, and R. Ahlrichs, J. Chem. Phys. **97**, 2571 (1992).
- [399] W. H. Miller, N. C. Handy, and J. E. Adams, J. Chem. Phys. **72**, 99 (1980).
- [400] G. A. Natanson, Mol. Phys. **46**, 481 (1982).
- [401] P. Pulay and G. Fogarasi, J. Chem. Phys. **96**, 2856 (1992).
- [402] C. F. Jackels, Z. Gu, and D. G. Truhlar, J. Chem. Phys. **102**, 3188 (1995).
- [403] K. A. Nguyen, C. F. Jackels, and D. G. Truhlar, J. Chem. Phys. **104**, 6491 (1996).
- [404] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**, 8302 (1998).
- [405] M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).
- [406] V. Bakken and T. Helgaker, J. Chem. Phys. **117**, 9160 (2002).
- [407] J. E. Bright Wilson, J. Decius, and P. C. Cross. *Molecular Vibrations* (McGraw-Hill, New York, 1955).
- [408] M. Tinkham. *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
- [409] T. Hayashi, K. Tanaka, and M. Haruta, J. Catal. **178**, 566 (1998).
- [410] G. Mills, M. S. Gordon, and H. Metiu, J. Chem. Phys. **118**, 4198 (2003).
- [411] E. M. Fernández, J. M. Soler, and L. C. Balbás, Phys. Rev. B **73**, 235433 (2006).
- [412] L. Xiao, B. Tolberg, X. Hu, and L. Wang, J. Chem. Phys. **124**, 114309 (2006).

- [413] X.-B. Li, H.-Y. Wang, X.-D. Yang, Z.-H. Zhu, and Y.-J. Tang, J. Chem. Phys. **126**, 084505 (2007), and references therein.
- [414] B. Assadollahzadeh and P. Schwerdtfeger, J. Chem. Phys. **131**, 064306 (2009), and references therein.
- [415] L. M. Molina and B. Hammer, J. Chem. Phys. **123**, 161104 (2005).
- [416] J. L. Bentz, R. M. Olson, M. S. Gordon, M. W. Schmidt, and R. A. Kendall, Comp. Phys. Commun. **176**, 589 (2007).
- [417] R. M. Olson, J. L. Bentz, R. A. Kendall, M. W. Schmidt, and M. S. Gordon, J. Chem. Theory Comput. **3**, 1312 (2007).
- [418] W. J. Stevens, M. Krauss, H. Basch, and P. G. Jasien, Can. J. Chem. **70**, 612 (1992).
- [419] L. Xiao and L. Wang, Chem. Phys. Lett. **392**, 452 (2004).
- [420] G. Bravo-Pérez, I. Garzón, and O. Novaro, J. Mol. Struct.: THEOCHEM **493**, 225 (1999).
- [421] G. Bravo-Pérez, I. Garzón, and O. Novaro, Chem. Phys. Lett. **313**, 655 (1999).
- [422] E. M. Fernandez and L. C. Balbas, Phys. Chem. Chem. Phys. **13**, 20863 (2011).
- [423] P. Gruene, B. Butschke, J. Lyon, D. Rayner, and A. Fielicke, Z. Phys. Chem. **228**, 337 (2014).
- [424] P. Piecuch, M. Włoch, and A. J. C. Varandas, Theor. Chem. Acc. **120**, 59 (2008).
- [425] K. R. Yang, A. Jalan, W. H. Green, and D. G. Truhlar, J. Chem. Theory Comput. **9**, 418 (2013).
- [426] M. Schütz, J. Chem. Phys. **116**, 8772 (2002).
- [427] K. P. Huber and G. Herzberg. *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).

- [428] L. Bytautas, T. Nagata, M. S. Gordon, and K. Ruedenberg, J. Chem. Phys. **127**, 164317 (2007).
- [429] A. Karton, B. Ruscic, and J. M. Martin, J. Mol. Struct.: THEOCHEM **811**, 345 (2007).
- [430] J. F. Stanton, J. Chem. Phys. **115**, 10382 (2001).
- [431] A. Köhn and A. Tajti, J. Chem. Phys. **127**, 044105 (2007).
- [432] E. Epifanovsky and A. I. Krylov, Mol. Phys. **105**, 2515 (2007).
- [433] T. Ichino, J. Gauss, and J. F. Stanton, J. Chem. Phys. **130**, 174105 (2009).