OVERCOMING THREE THEORETICAL CHALLENGES IN HIGH ACCURACY AB INITIO ELECTRONIC STRUCTURE THEORY

by

DAVID BRANDON MAGERS

(Under the Direction of Wesley D. Allen)

ABSTRACT

Modern ab initio electronic theory provides an accurate method to compute molecular properties of subchemical accuracy. Despite recent advances and developments, many challenges arise that require special treatment for particular systems and properties. Here three theoretical challenges are addressed. First, the need of a high accuracy multireference method is discussed to properly describe radical-radical abstraction reactions in chemical combustion. Second, a warning is given against black-box computations by documenting anomalous imaginary frequencies in common planar and linear systems when post-Hartree-Fock methods are applied with insufficient basis sets. Third, an analysis is provided to quantify the qualitative concept of intramolecular dispersion by examining individual pair correlation energies.

INDEX WORDS: coupled cluster theory, multireference coupled cluster theory, Møller-Plesset theory, basis set extrapolation, focal point analysis

OVERCOMING THREE THEORETICAL CHALLENGES IN HIGH ACCURACY AB INITIO ELECTRONIC STRUCTURE THEORY

by

DAVID BRANDON MAGERS

B.S., Mississippi College, 2009

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

OVERCOMING THREE THEORETICAL CHALLENGES IN HIGH ACCURACY AB INITIO ELECTRONIC STRUCTURE THEORY

by

D. BRANDON MAGERS

Major Professor:	Wesley D. Allen
Committee:	Henry F. Schaefer III Gary E. Douberly

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2014

ACKNOWLEDGEMENTS

Foremost I would like to thank my parents and my brother. They have been supportive of my fascination with math, science, and computers since day one. My mother's constant prayer for me should not go unacknowledged. She supported me even when I was at my worst. My brother was a healthy distraction from the daily grind. His humor and empathy have always impressed me. My father's insight into my life as a grad student in computational chemistry is unmeasurable. Without him I would not be receiving this degree. He has been a mentor, an advisor, a teacher, and a counselor in numerous ways. I have followed his footsteps very closely in my life and I would not have it any other way.

I would like to thank my fiancé. The first two and half years of our relationship tailed my grad school experience, and she has held my hand in support on a daily basis. On days when I wanted to quit, to days when I was full of excitement, she was a constant source of strength even from two states away.

So many friends have supported me in my quest for a Ph.D. Thank you to my small-group friends here in Athens. Your prayer and support was a rock in the storm. To my friends I left in Clinton, your constant reminders that I need to come "home" showed your love like none other.

I have had many teachers who have inspired me to pursue a life in science. I thank my high school chemistry teacher Coach West, calculus teacher Mrs. Kyzar, and math team coach Mr. Richie. Also I thank my instructors from undergrad Dr. Gann, Dr. Floyd, and Dr. Bishop.

To my advisor, Professor Allen, thank you for allowing me to join your research group. Your knowledge of quantum mechanics and theoretical methods is unmatched, and I only hope that some of that has rubbed off onto me. Thank you for the freedom to set my own schedule.

I thank my fellow CCQC members who have been here with me in this journey. Particularly I acknowledge Justin Turney, Jowa Wu, Alexander Sokolov, and Jay Agarwal for fruitful conversation and instruction.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 AB INITIO ELECTRONIC STRUCTURE THEORY

Over recent years, advances in modern quantum chemistry have made the ambition of subchemical accuracy possible. Today predicting thermochemical properties with accuracy better than 1.0 kcal mol^{-1} is possible through *ab initio* electronic structure theory, and thus has become an effective tool in chemical research. Computational methods provide a practical means for researchers to model reaction mechanisms, thermodynamic properties, kinetic rates, and other phenomena without ever entering the laboratory. These tools can save time and money, and in some cases predict properties that are currently unachievable in the laboratory with current techniques. Developing effective computational tools has challenges of its own such as maximizing computational power and optimizing software algorithms, while minimizing theoretical assumptions. The following review of electronic structure theory is by no means complete. The reader is encouraged to seek numerous resources for a more in-depth study.[1–4]

Modern quantum chemistry begins with its foundation in the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi \tag{1.1}$$

where Ψ is the wavefuction and $\hat{H}$ is the Hamiltonian operator. If there is no time-dependence in the Hamiltonian, the time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{1.2}$$

can be reduced from Eq 1.1. Any physical observable can be expressed mathematically by a Hermitian operator. For a given state, represented by the wave function, the solutions to the eigenvalue equation above are thus the observed energy. However solving this equation is by no means trivial, especially as the size of the system described by the wave function grows more complex. Solving the Schrödinger equation exactly is impractical, and thus the use of approximate wave functions and Hamiltonians is mandatory.

For a system of charged particles the nonrelativistic Hamiltonian in atomic units

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee} \tag{1.3}$$

or

$$\hat{H} = -\sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_i \frac{1}{2}\nabla_i^2 + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} - \sum_{Ai} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}} \tag{1.4}$$

is the sum of the kinetic and potential energy operators of the electrons and nuclei in the system. The commonly implemented Born-Oppenheimer, or "clamped nucleus," approximation assumes that since the nuclei are much more massive than the electrons, the nuclei are relatively stationary. This can be justified by classical time scales of the molecular motions. Classical electronic orbits exist on the attosecond time scale, while molecular rotations occur on the picosecond time scale and molecular vibrations occur on the femtosecond time scale. By

applying the Born-Oppenheimer approximation, the kinetic energy operator of the nuclei can be treated in a separate Schrödinger equation and the nuclear-nuclear repulsion term can be treated as a constant for the purpose for solving for the electronic wave function. This approximation yields the electronic Hamiltonian:

$$\hat{H}_e = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_{Ai} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}} + \hat{V}_{NN} \tag{1.5}$$

Even with this approximation, Eq. 1.2 is still difficult to solve exactly. The field of modern quantum chemistry is built around varying approximations to solutions of the electronic Schrödinger equation. A balance must be found between approximations, accuracy, and computational cost.

One of the biggest obstacles in solving the electronic Schrödinger equation is accounting for electron correlation in the system. Electrons, having a negative charge, repel each other. This instantaneous repulsion is difficult to compute. To achieve subchemical accuracy often over 99% of the total correlation needs to be recovered. For many-electron systems, the wave function depends on three spatial coordinates, r, and one spin coordinate ω. The wave function can be expanded in terms of antisymmetrized products of one-electron functions, or orbitals. A spin orbital is a function of both the spatial and spin coordinates of a single electron. A spatial orbital is a function of only the spatial coordinates. Spin orbitals can be written as a product of spatial orbitals and a spin function corresponding to either α or β spin. Discussion of construction of these spatial orbitals is reviewed in further detail in the next section. One might think that the wave function written in lowest approximation is a product of spin orbitals, each a function of the individual electrons. This is known as the Hartree Product. There are two major

issues with utilizing a Hartree Product as the wave function. In this approximation, the electrons are assumed not to interact with one another and the total wave function is just a product of one-electron wave functions. Physically, this is not the case. Secondly, the Hartree product does not obey the Pauli principle, which states that a wave function of fermions must be antisymmetric with respect to the interchange of any two particles. The Pauli principle gives rise to the Pauli Exclusion Principle, which states that no two identical fermions can occupy the same quantum state simultaneously. A solution to satisfy the Pauli principle is to write the wave function as a Slater determinant given in the general form as

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{1.6}$$

where χ is a spin orbital and $\mathbf{x}$ represents the three spatial coordinates and one spin coordinate of a single electron applied to a system with N electrons. Constructing a wave function as a Slater determinant means that every individual electron is associated with every orbital, as can be seen above. By including all possible Slater determinates, the exact solution to the Schrödinger equation can be found within the space spanned by the spin orbitals.

1.2 GAUSSIAN BASIS SETS

The one-electron functions needed to expand the wave function, Ψ, are often referred to as molecular orbitals. The molecular orbitals are expanded in a basis set, usually atomic orbitals

centered on each of the nuclei in the system. Particularly, Slater-type orbitals (STOs) describe atomic orbitals very well. These hydrogen-atom-like Slater functions

$$\phi_{abc}^{STO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r} \tag{1.7}$$

decay exponentially at long range. N is the normalization constant, a, b, and c dictate the angular momentum, ζ controls the size of the orbital, and r is the distance from the nucleus. The use of Slater functions is very accurate; however, practically these functions are difficult to integrate computationally.

In order to approximate Slater-type orbitals, a linear combination of Gaussian-type orbitals can be employed.

$$\phi_{abc}^{GTO}(x,y,z) = Nx^a y^b z^c e^{-\zeta r^2} \tag{1.8}$$

Because they are easier to integrate, Guassian-type orbitals reduce the computational cost, even though more of them are needed to describe the molecular orbitals. The same notation is used for Eq. 1.8 as was given in Eq. 1.7. Unlike Slater functions, Gaussian functions have a zero slope at $r = 0$ and decay much more quickly at long range. Most of the popular basis sets used today employ Gaussian-type orbitals.

Though there are hundreds of basis sets available in the literature, a few stand out for being popular, but with good reason. Notably the basis sets constructed by Dunning and coworkers[5–10] are designed specifically for post-Hartree-Fock correlation methods. Hartree-Fock theory and other methods are described in further detail in following sections. Dunning basis

sets are denoted generally as cc-pV*X*Z (where *X* = D, T, Q, 5, 6…), which stands for correlation-consistent, polarized valence, X-zeta basis set. Dunning basis sets are employed in the research described in Chapters 2, 3, and 4 and thus an in-depth look at their construction is warranted. Dunning basis sets are constructed to converge smoothly to the complete basis set limit with increasing cardinal number, *X*. In over-simplified terms, larger Dunning basis sets give better results. This aspect is extremely useful for a computational chemist's goal to achieve sub-chemical accuracy. By monitoring the convergence of increasing the basis set size, one can deduce the error due to utilizing an incomplete basis set. Also, clever extrapolation schemes and analyses are available to predict the complete basis set limit. One such analysis is described in detail in section 1.6 of this chapter.

Standard Dunning basis sets all include polarization functions. Polarization adds some needed flexibility to the orbitals. In a bonding environment, often a nearby atom's orbitals might distort another orbital. This property is described by mixing atomic orbitals with orbitals of higher angular momentum. For example, an s orbital can be mixed with a p orbital, or a p orbital with a d orbital.

Often Dunning basis sets are altered to better describe particular systems or situations. One of these includes adding diffuse functions. These augmented Dunning basis sets, denoted aug-cc-pV*X*Z,[6] add one additional diffuse function for each angular momentum shell currently present in the basis set. Diffuse functions extend further away from the nucleus, and therefore describe systems with lone pairs, anions, or systems with important long range interactions such as van der Waals complexes. Besides adding additional diffuse functions, often times the correlation energy contributed by the core electrons plays a non-negligible role. Standard Dunning basis are constructed to leave the core electrons uncorrelated. In order to do an all-

electron computation, the cc-pCV*X*Z Dunning basis sets[10] need to be used. These basis sets are designed to correlate the core electrons in post-Hartree-Fock methods.

The Dunning series of basis sets are ideal for many correlation methods and are utilized in the work described in chapters 2, 3, and 4. The use of well-constructed basis sets has helped improve the quality of electronic structure theory substantially.

1.3 HARTREE-FOCK THEORY

Hatree-Fock (HF) theory is a foundation in ab initio electronic structure methods. Hartree-Fock is the basis for molecular orbital theory and the starting place for many theories that approximate the electron correlation. At its core, HF theory is a first approach to solving the Schrödinger equation, though these approximations degrade the accuracy substantially. Hartree-Fock assumes that the wave function can be represented by a single Slater determinant, which falls short of the exact solution. Also HF does not include electron correlation, which is the most substantial downfall of the theory. The electron-electron instantaneous repulsion is ignored, and instead each electron feels the average Coulomb potential from all the other electrons together; thus, is called a mean field theory. However, in order to achieve subchemical accuracy, electron correlation is crucial and must be included. Despite its shortcomings, today Hartree-Fock is an inexpensive method and an excellent starting place for many other theories. Hartree-Fock is used as a starting point for theories utilized in the research in Chapters 2, 3, and 4.

Hartree-Fock assumes the Born-Oppenheimer approximation and utilizes the electronic Hamiltonian given in Eq. 1.5. Commonly this expression can be simplified using a new notation

$$\widehat{H}_{el} = \sum_i h(i) + \sum_{i<j} v(i,j) + V_{NN} \tag{1.9}$$

where the first term is the one-electron operator and the second term is the two-electron operator. As a reminder, the last term, the nuclear-nuclear repulsion potential energy, is only a constant in the Born-Oppenheimer approximation and does not affect the eigenfunctions. At this point the variational theorem is applied, which states that the expectation value of $\widehat{H}_e$ over a normalized trail wave function is greater than or equal to the ground-state energy ε_0.

$$E_{el} = <\Psi|\widehat{H}_{el}|\Psi> = \int_{-\infty}^{\infty} d\mathbf{r}\Psi^*\widehat{H}_{el}\Psi \geq \varepsilon_0 \tag{1.10}$$

Because of this theorem, the orbitals can be varied to minimize the energy within the space spanned by the orbitals. The Hartree-Fock energy can be expressed in terms of one- and two-electron integrals.

$$\begin{gathered} E_{HF} = \sum_i <i|h|i> + \frac{1}{2}\sum_{ij}[ij|jj] - [ij|ij] \\ \text{where } <i|h|i> = \int d\mathbf{x}_1 \chi_i^*(\mathbf{x}_1) h(\mathbf{r}_1) \chi_j(\mathbf{x}_1) \\ \text{and } [ij|kl] = \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_i^*(\mathbf{x}_1)\chi_j(\mathbf{x}_1)\frac{1}{r_{12}}\chi_k^*(\mathbf{x}_2)\chi_l(\mathbf{x}_2) \end{gathered} \tag{1.11}$$

Here χ_i denotes molecular orbitals. Some initial orbitals must be given and optimized through an iterative process. The Hartree-Fock equations for the orbitals are

$$f(\mathbf{x}_1)\chi_i(\mathbf{x}_1) = \epsilon_i\chi_i(\mathbf{x}_1)$$
$$\text{where } f(\mathbf{x}_1) = h(\mathbf{x}_1) + \sum_j J_j(\mathbf{x}_1) - K_j(\mathbf{x}_1) \tag{1.12}$$

and *f* is the Fock operator, *J* is the Coulomb operator, and *K* is the exchange operator.

The Eq. 1.12 above is difficult to solve, so a basis set is introduced. Here the molecular orbitals are formed by a linear combination of atomic orbitals

$$\chi_i = \sum_{\mu=1}^{K} C_{\mu i}\tilde{\chi}_\mu \tag{1.13}$$

where $\tilde{\chi}_\mu$ denotes the atomic orbital basis functions and $C_{\mu i}$ is the molecular-orbital coefficients. Substituting Eq. 1.13 into Eq. 1.12 results in the Hartree-Fock-Roothaan equations.[11] The result is a very solvable equation given as

$$\sum_\nu F_{\mu\nu}C_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu}C_{\nu i} \tag{1.14}$$

simplifying some of the notation into what is called the overlap matrix, **S**, and the Fock matrix, **F**. **F** must be diagonalized to obtain the molecular orbital coefficients, **C**, and the orbital energies, ϵ. The matrix representation of Eq. 1.14 must be solved iteratively since the Fock matrix depends on the orbital coefficients and vice versa. For this reason, the Hatree-Fock-Roothaan equations are called a self-consistent-field method.

1.4 MØLLER-PLESSET THEORY

As described above, Hartree-Fock theory does not account for electron correlation, the instantaneous electron-electron repulsion. This electron correlation is vital to describe many chemical systems and properties, and thus many post-Hatree-Fock correlation methods have been developed to improve the solution to the Schrödinger equation. One of these methods is Møller-Plesset perturbation theory.[3] This theory begins with a Hartree-Fock computation, and then improves the result by treating the difference between the exact and HF Hamiltonian as a perturbation. Unlike Hartree-Fock, Møller-Plesset Perturbation is not variational, and thus is not an upper bound to the exact energy.

Møller-Plesset perturbation theory begins with an unperturbed reference Hamiltonian then adds a perturbation.

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} \tag{1.15}$$

$\hat{H}_0$ is the reference Hamiltonian, λ is an ordering parameter generally set to one, and $\hat{V}$ is the perturbation operator. The perturbation operator is the difference between the instantaneous electron-electron interaction and the Hartree-Fock mean field approximation.

$$\hat{V} = \sum_{i<j} \frac{1}{r_{ij}} - \sum_{i} v^{HF}(i) \tag{1.16}$$

The eigenfunctions and eigenvalues can be expanded in a power series involving the order of perturbations. Optimally this expansion series will converge, and thus can be truncated.

Unfortunately this is not always the case. Commonly the expansion is truncated to the second-order term to give second-order Møller-Plesset perturbation theory. The formulation of the terms up through second-order is described below.

The reference Hamiltonian is taken as the sum of the Fock operators over all the electrons

$$\hat{H}_0 = \sum_i f(i) \tag{1.16}$$

and thus the zeroth-order energy, $E_0^{(0)}$, is just a sum of the Hartree-Fock orbital energies. In a closed-shell case, this sum double counts the electron-electron repulsion energy since each orbital is doubly occupied. The first-order correction is the expectation value of the perturbation operator, $\hat{V}$, over the reference state, the Hartree-Fock Slater determinant.

$$E_0^{(1)} = < \Psi_0 | \hat{V} | \Psi_0 > = -\frac{1}{2} \sum_{ab} < ab||ab > \tag{1.17}$$

The sum of the zeroth- and first-order correction is, thus, the Hartree-Fock energy. The first correction to the Hartree-Fock energy does not arise until the second-order perturbation correction is included.

The second-order correction is

$$E_0^{(2)} = \sum_{\mu \neq 0} \frac{| < \Psi_0^{(0)} | \hat{V} | \Psi_\mu^{(0)} > |^2}{E_0^{(0)} - E_\mu^{(0)}} \tag{1.18}$$

where $\Psi_{\mu}^{(0)}$denotes Slater determinants formed from excitations out of the reference wave function. Only double excitations will contribute to the energy. According to Brillouin's theorem,[3] singly excited determinants do not directly interact with the reference wave function. Triple and higher order excitations will not contribute because the perturbation is a two-particle operator. The term $E_{\mu}^{(0)}$ is simply the zeroth-order energy of the doubly excited determinant found via $\widehat{H}_0|\Psi_{ij}^{ab}>$. The final second-order Møller-Plesset perturbation energy is given as $E_{MP2} = E_{HF} + E_0^{(2)}$.

Møller-Plesset perturbation theory provides a decent estimate of dynamical electron correlation improving on Hartree-Fock. The computational cost is moderate, nominally scaling as N^5 with system size.

1.5 COUPLED CLUSTER THEORY

As modern quantum chemistry advanced, a huge leap forward was made with the development and implementation of coupled cluster theory.[12–14] Today coupled cluster theory is widely accepted as one of the most accurate ab initio methods to date. A key aspect to the quality of this theory is that in the full expansion including all possible excitations, coupled cluster is equivalent to Full Configuration Interaction which is the exact solution to the Schrödinger equation in the space spanned by the basis set.

The coupled cluster wave function is formed via the exponential of an excitation operator, $\hat{T}$.

$$\left|\Psi_{CC}>=e^{\hat{T}}\right|\Psi_0> \tag{1.19}$$

Here Ψ_0 is a single Slater reference determinant typically computed by Hartree-Fock. The excitation operator can be written as a summation of different levels of excitation or particle-rank

$$\hat{T}=\hat{T}_1+\hat{T}_2+\hat{T}_3+\ldots \tag{1.20}$$

where the subscript denotes the number of excitations from the reference wave function. The result of these excitation operators on a single determinant forms singly, doubly, etc. excited determinants

$$\hat{T}_1|\Phi_0>=\sum_{i,a}t_i^a\Phi_i^a \tag{1.21}$$

$$\hat{T}_2|\Phi_0>=\sum_{i<j}\sum_{a<b}t_{ij}^{ab}\Phi_{ij}^{ab} \tag{1.22}$$

where $i, j, \ldots$ denote occupied spin orbitals and $a, b, \ldots$ denote virtual spin orbitals. The t's are cluster amplitudes that need to be determined. Employing a similarity transformation, the energy and amplitude equations are found by projection forumlas.

$$E_{CC}=<\Phi_0\left|e^{-\hat{T}}\hat{H}e^{\hat{T}}\right|\Phi_0> \tag{1.23}$$

$$0=<\Phi_{ij\ldots}^{ab\cdots}\left|e^{-\hat{T}}\hat{H}e^{\hat{T}}\right|\Phi_0> \tag{1.24}$$

As shown in Eq. 1.20, the $\hat{T}$ operator can be expanded to include the excitation of all the electrons in the system. Higher excitation determinants contribute less to the total electron correlation, and therefore can be excluded in many instances while still recovering the majority of the electron correlation. This truncation gives coupled cluster its notation. By including only single and double excitations ($\hat{T} = \hat{T}_1 + \hat{T}_2$), the notation is denoted CCSD where S stands for single and D stands for double excitations.

Two properties that benefit coupled cluster theory is being both size extensive and size consistent. If a method is size extensive, then the energy scales linearly in respect to the size of the system. If a method is size consistent then it correctly gives the energy of two non-interacting systems separated by distance. In other words, the energy of the system A + B must equal the energy of A plus the energy of B. Generally this is associated with the dissociation of a system into fragments. Since coupled cluster is based on a Hartree-Fock wave function, it is not necessarily size consistent for all systems.

As mentioned, increasing the number of excitations, and thus increasing the size of the $\hat{T}$ operator, improves the quality of the energy by recovering more of the total electron correlation. Unfortunately higher order excitations many times are computationally expensive. In order to reduce computational cost, many times higher excitation orders are approximated with a perturbative treatment. A very popular method is fully computing the coupled cluster single and doubles excitation energy and then perturbatively including the triples contribution. This theory is denoted CCSD(T). For a majority of systems and physical properties, CCSD(T) with a reasonable basis set will achieve subchemical accuracy.

1.6 FOCAL POINT ANALYSIS

As seen in the sections above, practical applications to solve the Schrödinger equation include layers of approximations, and thus mindful construction of molecular wave functions and basis sets. Unfortunately the computational cost to solve the Schrödinger equation exactly with full configuration interaction at a complete basis set is too great. However, as mentioned briefly above, Dunning basis sets are designed to systematically converge to the complete basis set limit. Also coupled cluster theory converges to the correlation limit, and thus the exact energy. By observing the convergence of systematically improving both the basis set and the correlation treatment, the feasibility of achieving these limits is improved greatly. The focal point analysis[15–19] (FPA) scheme employs such an approach to achieve the exact result.

The focal point analysis targets both the correlation limit as well as the basis set limit. As the number of excitations is increased in the coupled cluster equations the energy approaches the FCI energy. Coupled cluster theory is discussed in more detail in the previous section 1.5. The contribution to the energy from each additional excitation will naturally decrease. In order to observe the convergence to the correlation limit the following hierarchy is generally followed: HF, MP2, CCSD, CCSD(T), CCSDT, CCSDT(Q), etc. The convergence is easily observed by computing the increment (δ) between one level of theory and the next step up in the hierarchy. For instance, the CCSDT(Q) correction to the energy is the difference from the CCSDT energy:

$$\delta[\mathrm{CCSDT(Q)}] = \Delta E_{\mathrm{CCSDT(Q)}} - \Delta E_{\mathrm{CCSDT}} \tag{1.25}$$

As discussed in section 1.5, it is noteworthy that a perturbative treatment of the energy recovered by a term in the $\hat{T}$ operator is just an estimation of the full energy at that level.

The focal point analysis requires the use of a collection of basis sets, such as the Dunning basis sets, that systematically improve. In order to achieve the complete basis set limit, a variety of extrapolation schemes is utilized throughout the literature. A standard extrapolation scheme[20]

$$E_X^{HF} = A + Be^{-CX} \tag{1.26}$$

for the Hartree-Fock method uses a three-parameter exponential function. For post Hartree-Fock correlation methods, a different extrapolation scheme[21] is common. This extrapolation

$$E_X^{corr} = A + BX^{-3} \tag{1.27}$$

is a two-parameter function. For this formula, only the correlation energy is extrapolated and not the total energy. In a standard implementation of these equations, lower levels of theory, such as Hartree-Fock and MP2, are computed up through a large basis set, such as cc-pV5Z. In this example, the Hartree-Fock cc-pVXZ (X = T, Q, 5) energies can be extrapolated to the complete basis set limit with Eq. 1.26. While moving up the ladder, incorporating higher excitation levels in the coupled cluster computations, the energy increments are expected to converge to the exact answer sooner with increasing basis set size. In essence, at higher levels of theory smaller basis sets are used. This is especially true since the focal point table is written as an incremented correction as described by Eq. 1.25 and thus describes the correction to the one-step-lower level of theory instead of the total energy.

By monitoring the convergence of the coupled cluster correlation corrections with inclusion of higher levels of excitations, while also extrapolating to the complete basis set limit at lower levels of theory, the computational cost can be kept affordable while achieving high accuracy in estimating the exact energy. The focal point analysis is a powerful tool to aid the computational chemist to achieve the best answer possible. For example, instead of computing a CCSDTQP/cc-pV6Z energy by brute force, which is impossible except for the smallest of systems, the FPA approach provides a means to practically achieve high accuracy results.

1.7 RESEARCH PROSPECTUS

The previous sections provide a brief overview of *ab initio* electronic structure theory. As the reader has probably realized, a lot of approximations are utilized on multiple layers to achieve meaningful results. The casual scientist is warned against using computational chemistry as a "black box" tool. Understanding the mathematics and physics behind the computer screen is crucial. The appropriate level of theory, basis set, and analyses need to be used to answer a particular problem. There are many aspects of electronic structure theory that are difficult for a theoretical chemist to answer without a careful look at the problem. The following chapters in this dissertation address three different theoretical challenges for describing physical phenomena.

In Chapter 2, a multireference coupled cluster theory (Mk-MRCC) is utilized to describe potential energy surfaces of radical-radical hydrogen abstraction reactions which are important in combustion models. Mk-MRCC theory is described in detail in that chapter. For this type of reaction between two radicals, standard theoretical approaches fail miserably. The root of the problem comes from single-reference methods, such as coupled cluster, being unable to properly

describe two radicals at long range. This theoretical inability is overcome through the use of multireference methods.

In Chapter 3, planar and linear molecules are unphysically predicted to be non-planar and bent, respectively, by using improper basis sets. Specifically, the use of post-Hartree-Fock methods, such as coupled cluster, with Pople basis sets predict improper geometries for many planar and linear systems. Without knowledge of the correct implementation of these basis sets, a computational chemist could succumb to this pitfall. This theoretical challenge is solved by using medium to large correlation-consistent basis sets such as the Dunning basis sets.

In Chapter 4, the concept of quantifying intramolecular dispersion in 1-(1-diamantyl) diamantane is explored. Generally dispersion is thought of as an intermolecular property and theoretically thought of as a piece of the total correlation energy. Dispersion is not rigorously defined, but it is generally considered to exist in the long-range portion of electron-electron correlation interactions. In order to quantify this property, an analysis of the correlation contribution from individual electron pairs was utilized to categorize the total correlation energy to either dispersion or non-dispersion. This method allowed an abstract concept to be quantified with *ab initio* methods.

1.8 REFERENCES

(1) McQuarrie, D. A.; Simon, J. D. *Physical Chemistry A Molecular Approach*; University Science Book: Sausalito, CA, 1997.

(2) Griffiths, D. J. *Introduction to Quantum Mechanics*; Second.; Pearson Prentice Hall: Upper Saddle River, NJ, 2005.

(3) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, NY, 1996.

(4) Sakurai, J. J. *Modern Quantum Mechanics*; Revised.; Addison-Wesley: Reading, MA, 1994.

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

(6) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(7) Woon, D. E.; Dunning Jr., T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(8) Woon, D. E.; Dunning Jr., T. H. *J. Chem. Phys.* **1994**, *100*, 2975.

(9) Woon, D. E.; Dunning Jr., T. H. *J. Chem. Phys.* **1995**, *103*, 4572.

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

(11) Roothaan, C. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

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

(13) Crawford, T. D.; Schaefer, H. F. *Rev. Comput. Chem.* **2000**, *14*, 33–136.

(14) Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

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

(16) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 9751.

(17) Császár, A. G.; Tarczay, G.; Leininger, M. L.; Polyansky, O. L.; Tennyson, J.; Allen, W. D. In *Spectroscopy from Space*; Demaison, J.; Sarka, K.; Cohen, E., Eds.; NATO Science Series; Springer Netherlands, 2001; Vol. 20, pp. 317–339.

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

(19) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2004**, *120*, 11586–11599.

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

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

CHAPTER 2

REACTION PROFILES FOR RADICAL-RADICAL HYDROGEN ABSTRACTION VIA MULTIREFERENCE COUPLED CLUSTER THEORY [†]

[†] Chia-Hua Wu,[‡] D. Brandon Magers,[‡] Lawrence B. Harding, Stephen J. Klippenstein, and Wesley D. Allen. To be submitted to *Journal of Chemical Theory and Computation*.

[‡] These authors contributed equally.

2.1 ABSTRACT

Radical-radical abstractions in hydrocarbon oxidation chemistry are disproportionation reactions that are generally exothermic with little or no barrier, yet are underappreciated and poorly studied. Such challenging multireference electronic structure problems are tackled here using the recently developed state-specific multireference coupled cluster methods Mk-MRCCSD and Mk-MRCCSD(T), as well as the companion perturbation theory Mk-MRPT2 and the popular MRCISD, MRCISD+Q, and CASPT2 approaches. Reaction paths are investigated for five prototypes involving radical-radical hydrogen abstraction: $H + BeH \rightarrow H_2 + Be$, $H + NH_2 \rightarrow H_2 + NH$, $CH_3 + C_2H_5 \rightarrow CH_4 + C_2H_4$, $H + C_2H_5 \rightarrow H_2 + C_2H_4$, and $H + HCO \rightarrow H_2 + CO$. Full configuration interaction (FCI) benchmark computations for the H + BeH, $H + NH_2$, and H + HCO reactions prove that the Mk-MRCCSD(T) potential energy curves display superior accuracy, within mean absolute errors of only 0.2 kcal mol^{-1}. To facilitate studies of combustion kinetics, energetics for the $CH_3 + C_2H_5$, $H + C_2H_5$, and H + HCO reactions were computed at each level of theory with correlation-consistent basis sets (cc-pV*X*Z, *X* = T, Q, 5) and extrapolated to the complete basis set (CBS) limit. The rigorous Mk-MRCCSD(T)/CBS results demonstrate unequivocally that these three reactions proceed with no barrier in the entrance channel, contrary to some earlier predictions. Mk-MRCCSD(T) also reveals that the economical CASPT2 method performs well for large interfragment separations but may deteriorate substantially at shorter distances.

2.2 INTRODUCTION

While the role of radical-radical recombination in hydrocarbon combustion is widely appreciated,[1-5] radical-radical hydrogen abstraction has not received much attention. As shown generically in the following equation, "when one of a pair of encountering radicals has a hydrogen atom β to the radical center, two stable molecules are produced when this hydrogen atom is abstracted by the other radical:"[6]

$$R\bullet + HXY\bullet \rightarrow R\cdots H\cdots XY \rightarrow RH + XY$$

Such disproportionation reactions are expected to be highly exothermic and devoid of large activation barriers.

Several studies[2, 7-10] have emphasized that disproportionation is a viable radical termination reaction. Early kinetic models argued that alkyl radical disproportionation can occur when the radicals approach within ~4 Å.[2, 7] Laser photolysis-kinetic spectroscopy studies found that disproportionation is competitive with recombination in vinyl-vinyl and vinyl-methyl reactions.[8-10] The disproportionation of hydroxyl radicals, OH + OH $\rightarrow$ H_2O + O, has been the subject of copious experimental[11-15] and theoretical[16-21] research; however, recent values for the rate constant of this reaction have differed by a factor of ~2.[11, 12, 15] Relevant to the formation of polycyclic aromatic hydocarbons (PAHs) in combustion environments, phenyl radicals were observed in shock tube experiments[22] to not only recombine but also disproportionate to benzene + benzyne. An accompanying theoretical kinetic analysis[22] showed comparable rates for the recombination and disproportionation of phenyl radicals in the temperature range 1000–2000 K.

Recently discovered "roaming radical" reactions can occur when a covalent bond in an energized molecule cleaves homolytically but the two resulting radical fragments remain loosely

bound and do not fully separate. Roaming on the flat, long-range potential energy surface allows one radical fragment to abstract a hydrogen atom from the backside of the other to form two stable molecular products.[23-29] Fundamentally this intramolecular abstraction process is the same as the disproportionation reaction outlined above. The roaming radical mechanism was first used to explain the existence of a second molecular channel of formaldehyde dissociation to yield vibrationally "hot" H_2 and rotationally "cold" CO, in contrast to the highly rotationally excited CO produced from the conventional unimolecular TS.[29] A similar mechanism was also found in the photodissociation of acetaldehyde[24, 25] (methyl + formyl radicals) and dimethyl ether (methyl + methoxy radicals).[26] In the decomposition of alkanes, the roaming radical mechanism generates a significant fraction of non-H-atom products.[23, 26-28] Both shock tube experiments and *ab initio* transition state theory computations show that in the dissociation of propane, 10% of the products arise from the roaming radical disproportionation $CH_3 \cdots C_2H_5 \rightarrow CH_4 + C_2H_4$.[28] The roaming radical mechanism is also the dominant pathway for isobutane and neopentane decomposition at high temperatures (1200-1500 K).[27]

Prototypically, radical-radical hydrogen abstraction begins with two open-shell reactants and ends with two closed-shell products, and thus the electronic transformation is intrinsically multireference in nature.[4] An unsound treatment of this challenging electronic structure problem can yield incorrect results. For instance, using single-reference CCSD(F12) theory for the phenyl + phenyl system gives disproportionation rates one order of magnitude lower than the corresponding recombination rates in the temperature range 1000–2000 K.[30] In contrast, multireference CASPT2(2,2) computations extrapolated to the CBS limit yield comparable disproportionation and recombination rates. [22]

A recent quantum chemical investigation[31] found the methyl + ethyl disproportionation reaction to be barrierless with the CR-CC(2,3) and MCQDPT(4,4) methods, whereas CAS(4,4) gave a barrier of ~5 kcal mol^{-1}. Previous G2M(CC2)//B3LYP/6-311+G(3*df*,2*p*)[32] and QCISD/6-311+G(2*d*, 2*p*)//MP2/6-311+G (2*d*, 2*p*)[33] computations found that the two radicals first form a long-range complex before encountering an energetic barrier of 1.5 – 6 kcal mol^{-1} in the hydrogen abstraction step. Harding and Klippenstein[23, 34] used the CASPT2/aug-cc-pVDZ method to map out potential energy surfaces (PESs) for roaming radical kinetics in the systems dimethyl ether and acetaldehyde. They reported that the computed reaction rate becomes significantly more accurate when a one-dimensional PES correction is used to improve the reaction path energetics given by CASPT2/aug-cc-pVDZ.[23, 26, 34]

Overall, the previous work on both disproportionation and roaming radical reactions highlights the need for definitive multireference electronic structure studies to elucidate radical-radical hydrogen abstractions. For this purpose, we have applied our state-of-the-art multireference coupled cluster method (Mk-MRCC), as well as conventional MRCISD, MRCISD+Q, CASPT2, and FCI techniques, to carefully investigate five prototype reactions:

$$H + BeH\ (^2\Sigma^+) \rightarrow H_2 + Be \quad (1)$$

$$H + NH_2\ (^2B_1) \rightarrow H_2 + NH\ (^1\Delta) \quad (2)$$

$$CH_3 + C_2H_5\ (^2A') \rightarrow CH_4 + C_2H_4 \quad (3)$$

$$H + C_2H_5\ (^2A') \rightarrow H_2 + C_2H_4 \quad (4)$$

$$H + HCO\ (^2A') \rightarrow H_2 + CO \quad (5)$$

State-specific Mk-MRCC multireference coupled-cluster theory, originally introduced by Mahapatra et al.,[35, 36] has been recently developed into a practical method for chemical applications.[37-42] The Mk-MRCC wave function $\left|\Psi_\alpha\right\rangle$ of state α is constructed using the Jeziorski-Monkhorst ansatz,[43] which involves a set of *d* reference determinants $\left|\Phi_\mu\right\rangle$ within some active space acted on by individual exponential operators $e^{\hat{T}^\mu_\alpha}$ and multiplied by configuration interaction coefficients c^μ_α:

$$\left|\Psi_\alpha\right\rangle = \sum_{\mu=1}^{d} c^\alpha_\mu e^{\hat{T}^\mu} \left|\Phi_\mu\right\rangle \tag{6}$$

As usual, the cluster operators $\hat{T}^\mu_\alpha$ are truncated at a certain excitation level *n*. Mk-MRCC energies E_α and CI coefficients c^μ_α are determined as eigenvalues and eigenvectors of an effective Hamiltonian matrix,[38]

$$\sum_{v=1}^{d} H^{eff}_{\mu v} c^\alpha_v = E^\alpha c^\alpha_\mu \tag{7}$$

$$\text{where } H^{eff}_{\mu v} = \left\langle \Phi_\mu \middle| \hat{H} e^{\hat{T}^\mu} \middle| \Phi_v \right\rangle . \tag{8}$$

Sufficiency conditions are applied in Mk-MRCC theory that yield the equations for the cluster amplitudes,

$$\Delta_{ij..}^{ab..}(\mu) = \left\langle \Phi_{ij...}^{ab...}(\mu) \middle| \bar{H}_{\mu} \middle| \Phi_{\mu} \right\rangle c_{\mu} + \sum_{\nu(\neq\mu)} \left\langle \Phi_{ij...}^{ab...}(\mu) \middle| e^{-\hat{T}^{\mu}} e^{\hat{T}^{\nu}} \middle| \Phi_{\mu} \right\rangle H_{\mu\nu}^{eff} c_{\nu} = 0 \tag{9}$$

in which $\left|\Phi_{ij...}^{ab...}(\mu)\right\rangle$ is a determinant obtained from $\left|\Phi_{\mu}\right\rangle$ by a collective excitation from occupied orbitals (*ij*...) to virtual orbitals (*ab*...) and

$$\bar{H}_{\nu} = e^{-\hat{T}_{\alpha}^{\nu}} \hat{H}\ e^{\hat{T}_{\alpha}^{\nu}} \tag{10}$$

is the familiar coupled-cluster similarity-transformed Hamiltonian for a specific reference ν.

State-specific Mk-MRCC theory has the advantage of size extensivity, orbital invariance except in the active space, and avoidance of the intruder state problem.[38] Moreover, the coupling terms $\left\langle \Phi_{ij...}^{ab...}(\mu) \middle| e^{-\hat{T}_{\alpha}^{\mu}} e^{\hat{T}_{\alpha}^{\nu}} \middle| \Phi_{\mu} \right\rangle$ in eq 9 can be evaluated by a simple closed-form expression.[38] Mk-MRCC theory is one among numerous state-specific multireference coupled cluster approaches; a recent review nicely compares and assesses these different methods.[44] Capabilities are available for incorporating triple excitations into Mk-MRCC,[40] especially by means of a perturbative triples correction that yields a highly accurate Mk-MRCCSD(T) theory.[39] The number of successful chemical applications of Mk-MRCC theory is growing.[45-52] A companion multireference perturbation theory (Mk-MRPT2) has also been formulated and implemented,[41] but its accuracy has not been extensively tested or compared with related methods such as CASPT2.[53, 54]

2.3 METHODS

2.3.1 ELECTRONIC STRUCTURE THEORIES

In order to benchmark the Mk-MRCC performance in radical-radical abstraction processes, we have selected several prototypes (reactions (1) – (5)) for investigation. For each reaction, one-dimensional reaction paths were first generated by CASPT2 (second–order complete active space perturbation theory) constrained geometry optimizations using the reaction coordinate $S = R_1 - R_2$, as described in Figure 2.1. For each value of S at intervals of about 0.2 Å, all other internal coordinates were optimized. Our preliminary computations suggest that a simple CAS(2,2)[55, 56] reference wave function involving a 2-electron-in-2-orbital active space correctly describes the reaction in zeroth order, as discussed in the next section. Therefore, we focus here on multireference computations with this active space.

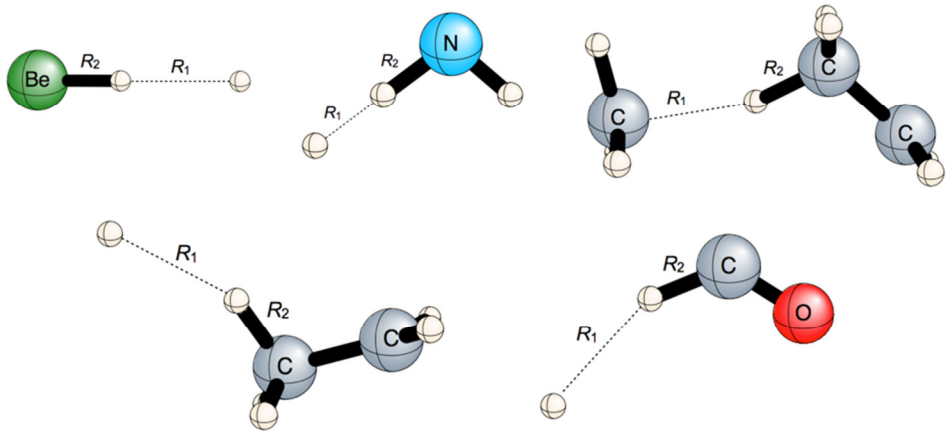


FIGURE 2.1: Radical-radical hydrogen abstraction reactions of eqs 1-5, as described by the reaction coordinate $S = R_1 - R_2$. Separated (reactants, products) correspond to $S = (+\infty,-\infty)$.

Several established multireference methods [MRCISD,[57, 58] MRCISD+Q (Davidson correction),[59] CASPT2[53, 54, 60]] were compared to the Mk-MRPT2,[41] Mk-MRCCSD,[37, 38] and Mk-MRCCSD(T)[39] approaches for computing the energy profiles. The MRCI computations were internally contracted as developed by Werner and Knowles.[57, 58] All MRCI and CASPT2 computations adopted the CAS(2,2) natural orbitals for the reference wave function. Single reference CCSD(T)[61, 62] and CCSDT(Q)[63, 64] computations were performed to show that even high-accuracy single reference approaches gradually break down in the multireference region of the reaction paths. Core electrons were frozen in all of electron correlation methods employed. For benchmark purposes, full configuration interaction[65-68] (FCI) energies were computed with small basis sets for reaction 1, 2, and 5. For the H + HCO reaction, the cc-pVDZ(*sp*) basis set was employed in which heavy-atom *d* functions and hydrogen *p* functions were removed from the standard cc-pVDZ[69] basis set. For the H + NH_2 reaction, the Pople 6-311G* basis set[70, 71] was used to compute the FCI energies. The electronic structure methods and basis sets applied in each reaction are listed in Table 2.1.

In order to facilitate studies of reaction kinetics, energies for reactions 3, 4, and 5 computed with correlation-consistent basis sets (cc-pV*X*Z)[69] were extrapolated to the complete basis set (CBS) limit.[72-76] Our analysis shows that the effect of core correlation on the energy profile of the entrance channel is quite limited;[77] hence, only valence electron correlation was treated in this study. In our energy extrapolation scheme, a three-parameter exponential function[78, 79] was used for the reference CASSCF extrapolations,

$$E_X^{\mathrm{CASSCF}} = E_{CBS}^{\mathrm{CASSCF}} + ae^{-bX} \quad (X = 3,4,5) \tag{11}$$

while the multireference electron correlation energies (ε) were extrapolated using the two-parameter formula[80]

$$\varepsilon_X = \varepsilon_{CBS} + bX^{-3} \quad (X = 3,4) \tag{12}$$

In addition, the geometries of both products and radical fragments for reaction 3, 4 and 5 were optimized at the single-reference CCSD(T)/cc-pCVQZ[69] level of theory with the ZPVE correction at CCSD(T)/cc-pVTZ. The overall energies of reaction were probed using CCSD(T) with the same energy extrapolation approach in eq 11 and 12. The Mk-MRPT2, Mk-MRCCSD and Mk-MRCCSD(T) computations were performed using the PSIMRCC routine[81] in the quantum chemical program Psi 3.4 version[82] as well as Psi4.[83] CCSDT(Q) computations were carried out using the MRCC program.[63] The MRCI, CASPT2, single reference CCSD(T) and full CI energies were obtained using the Molpro 2010.1 package.[84] The default options of the MRCI algorithm in Molpro were kept throughout the computations, while the level shift parameter[85] ε = 0.2 was imposed in CASPT2 iterations to prevent intruder state problems. The vibrational frequencies along the reaction path for reactions 3, 4, and 5 were computed using finite differences of analytical gradient. The reaction mode was projected out using the DRP procedure of Allen et al.[86, 87] The Fortran program INTDER[12] was used for the projected frequency analysis. The results of vibrational analyses along each reaction path are placed in the supporting information.

TABLE 2.1: Theoretical methods employed for radical-radical hydrogen abstraction reactions

Reaction	Reaction path geometry and frequencies	Energy profiles
$H + BeH \rightarrow H_2 + Be$	CASPT2/cc-pVTZ	FCI, Mk-MRCCSD(T), Mk-MRCCSD, Mk-MRPT2, MRCISD, MRCISD +Q, CASPT2, CCSDT(Q) Basis set: cc-pVTZ
$H + NH_2 \rightarrow H_2 + NH$	CASPT2/6-311G*	FCI, Mk-MRCCSD(T), Mk-MRCCSD, Mk-MRPT2, MRCISD, MRCISD +Q, CASPT2 Basis set: 6-311G*
$CH_3 + C_2H_5 \rightarrow CH_4 + C_2H_4$	CASPT2/cc-pVQZ	Mk-MRCCSD(T), Mk-MRCCSD, MRCISD, MRCISD+Q, CASPT2, CCSD(T) Basis sets: cc-pVTZ, cc-pVQZ, and CBS limit
$H + C_2H_5 \rightarrow H_2 + C_2H_4$	CASPT2/cc-pVQZ	Mk-MRCCSD(T), Mk-MRCCSD, Mk-MRPT2, MRCISD, MRCISD+Q, CASPT2, CCSD(T) Basis sets: cc-pVTZ, cc-pVQZ, and CBS limit
$H + HCO \rightarrow H_2 + CO$	CASPT2/cc-pVDZ(*sp*)	FCI, Mk-MRCCSD(T), Mk-MRCCSD, Mk-MRPT2, MRCISD, MRCISD+Q, CASPT2, CCSDT(Q) Basis set: cc-pVDZ(*sp*)
	CASPT2/cc-pVQZ	Mk-MRCCSD(T), Mk-MRCCSD, Mk-MRPT2, MRCISD, MRCISD+Q, CASPT2, CCSD(T) Basis sets: cc-pVTZ, cc-pVQZ, and CBS limit

2.3.2 REERENCE WAVE FUNCTION FOR MULTIREFERENCE TREATMENT

To diagnose the multireference problem, preliminary CAS(4,4) and CAS(2,2) computations were performed in the entrance channel of the methyl + ethyl reaction. Table 2.2 compares the leading CI coefficients in the CAS wave functions. The C_1 and C_2 coefficients in the CAS(4,4) wave function are very similar to their CAS(2,2) counterparts, while the CAS(4,4) C_3 and C_4 coefficients are quite small. Our analysis covers a wide range of reaction coordinate values (S = –3.0 ~ 3.0 Å) to ensure that the CAS(2,2) reference wave function is sufficient for dynamically-correlated Mk-MRCC computations along the entire reaction path. Similar conclusions are drawn for the other reactions in this study, and the results are provided in the supporting information.

In earlier Mk-MRCC studies, Evangelista et al.[38] found that in the simple F· + F· radical recombination, the Mk-MRCCSD results are more accurate when localized active MOs are used. Our preliminary analysis and Das et al.[88] share the same observation.[38] Therefore, in this work, localized orbitals were constructed from the active CAS (2,2) natural orbitals by the Boys orbital localization method.[89] This localization scheme was applied for the Mk-MRCC wave functions throughout the reaction path. Near the asymptotic limit of the exit channel, the two active orbitals become localized on one of the separated products. In reaction 1, as the hydrogen abstraction is near complete, the two active electrons will be localized in 2s and 3s orbitals of the beryllium atom. Similarly, in reactions 3, 4 and 5, the π and π* orbitals in the C_2H_4 and CO products comprise the active space of Mk-MRCC computations in the asymptotic limit. Reaction 2 represents an open-shell singlet system; therefore, the Boys orbital localization scheme does not apply for this reaction.

TABLE 2.2: Leading CI coefficients (C_i) in the CASSCF/cc-pVTZ wave functions along the entrance channel of the methyl + ethyl hydrogen abstraction reaction

	CAS(2,2)		CAS(4,4)			
S (Å)	C_1	C_2	C_1	C_2	C_3	C_4
–3.0	0.978	–0.207	0.974	–0.206	–0.090	> 0.050
–1.8	0.978	–0.206	0.974	–0.206	–0.089	> 0.050
–0.6	0.978	–0.208	0.973	–0.206	–0.088	> 0.050
–0.2	0.976	–0.217	0.968	–0.213	–0.076	> 0.050
0.0	0.972	–0.236	0.960	–0.240	–0.065	–0.063
0.2	0.953	–0.303	0.941	–0.301	–0.074	0.074
0.6	0.866	–0.499	0.862	–0.490	–0.071	0.051
1.0	0.793	–0.609	0.792	–0.603	–0.063	> 0.050
1.4	0.758	–0.653	0.756	–0.648	–0.062	0.054
1.8	0.738	–0.675	0.735	–0.672	–0.061	0.056
2.2	0.726	–0.688	0.723	–0.684	–0.060	0.057
2.6	0.714	–0.700	0.711	–0.696	–0.060	0.058
3.0	0.713	–0.701	0.710	–0.698	–0.059	0.059

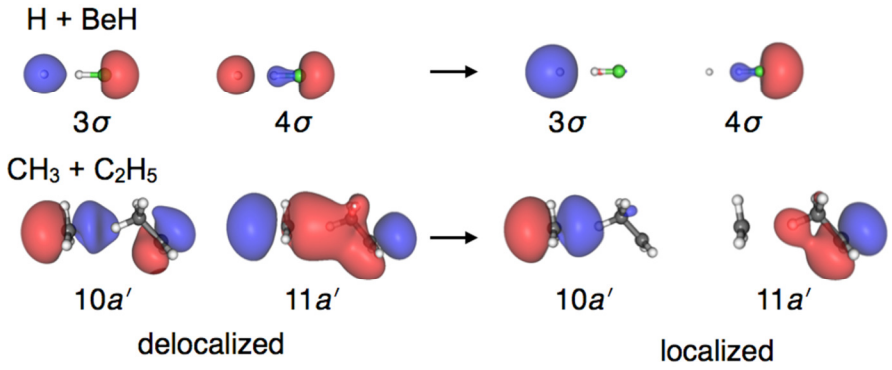


FIGURE 2.2: Boys localization of the active molecular orbitals of the H + BeH and $CH_3 + C_2H_5$ systems for S = 1.0 Å.

2.3.3 TIKHONOV REGULARIZATION IN MK-MRCC ITERATIONS

In the iterative procedure in Mk-MRCC theory, the cluster amplitudes are updated in each cycle using eq 13:[38]

$$t_{ij\ldots}^{ab\ldots}(\mu,\text{new}) = t_{ij\ldots}^{ab\ldots}(\mu,\text{old}) + \frac{\Delta_{ij\ldots}^{ab\ldots}(\mu)}{c_\mu^\alpha [f_{ii}(\mu) + f_{jj}(\mu) + \ldots - f_{aa}(\mu) - f_{bb}(\mu) + E^\alpha - H_{\mu\mu}^{eff}]} \tag{13}$$

The t-amplitude residuals $\Delta_{ij\ldots}^{ab\ldots}(\mu)$ are obtained from eq 9 and $f_{ij}(\mu)$ is the Fock matrix element for orbitals i and j. In eq 13, the CI coefficient c_μ^α appears in the denominator. Therefore, Mk-MRCC iterations could suffer from numerical instability when c_μ^α has a small value. Tikhonov regularization was employed to resolve this numerical issue.[90, 91] A small factor (ω) is introduced in c_μ^α to avoid the convergence issue, as shown in eq 14:[90]

$$\frac{1}{c_\mu^\alpha} \rightarrow \frac{c_\mu^\alpha}{[(c_\mu^\alpha)^2 + \omega^2]} \tag{4.14}$$

Inserting Eq. 14 into the cluster amplitude expression introduces a quadratical deviation from the exact Mk-MRCC energy when ω increases, as demonstrated in Figure 2.3 using the geometry at $S = 0.4$ for reaction 4. For all the Mk-MRCC energies we report, if numerical instability is observed in the Mk-MRCC iterations, several computations were then performed using different ω values and a quadratic extrapolation scheme is applied to obtain the correct energy ($\omega = 0$). The extrapolation scheme was tested for ten different geometries and the average error of the 3-point extrapolated Mk-MRCC energies is ~10^{-8} hartree.[92]

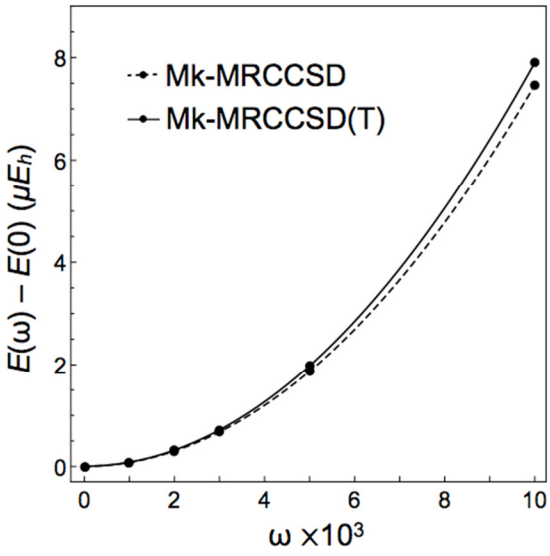


FIGURE 2.3: Quadratic dependence of cc-pVTZ Mk-MRCCSD and Mk-MRCCSD(T) electronic energies of the H + C_2H_5 system ($S = 0.4$ Å) on the Tikhonov parameter (ω). The reference values $E(0)$ of the Mk-MRCCSD and Mk-MRCCSD(T) energies are –79.4948772 and –79.5074588 E_h, respectively.

2.4 RESULTS AND DISCUSSION

2.4.1 $H + BeH \rightarrow H_2 + Be$

$H + BeH$ ($^2\Sigma^+$) is the simplest radical-radical abstraction reaction that produces closed-shell products ($H_2 + Be$). The BeH_2 model system[38-40, 93] has often served as a test case for multireference electronic structure methods. The reaction path was mapped out assuming that H approaches the BeH radical along the molecular axis, as shown in Figure 2.1.

The potential energy curves of the H + BeH reaction at different levels of theory are shown in Figure 2.4. The curve in best agreement with the FCI benchmark is CCSDT(Q). This is no surprise because in a four valence electron system with the core frozen, CCSDTQ is equivalent to FCI. Using the CAS(2,2) reference wave function, the MRCCSD(T) treatment shows accuracy superior to CASPT2, MRCISD, and MRCISD+Q. The Davidson correction improves the MRCISD prediction, but it overestimates the higher-order correlation correction, yielding a curve on the negative side of the FCI benchmark. Figure 2.5 shows the error in the binding energy relative to FCI at different levels of theory. Except in the PT2 cases, the errors only become substantial after the reaction commences in earnest ($S < 0.5$ Å). Among the multireference methods, Mk-MRCCSD(T) gives the smallest errors overall. The perturbation theories (CASPT2 and Mk-MRPT2) underestimate the binding energy significantly for this reaction. However, the (T) perturbative treatment clearly improves Mk-MRCCSD(T) over Mk-MRCCSD.

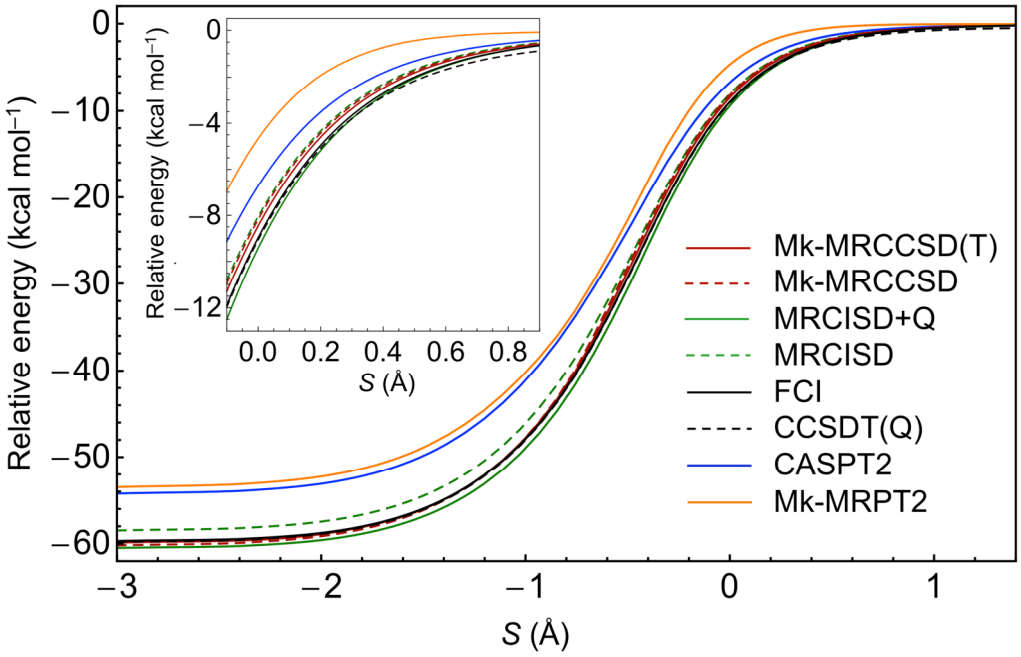


FIGURE 2.4: Potential energy curves for the entrance channel of H + BeH computed at numerous levels of theory with the cc-pVTZ basis set.

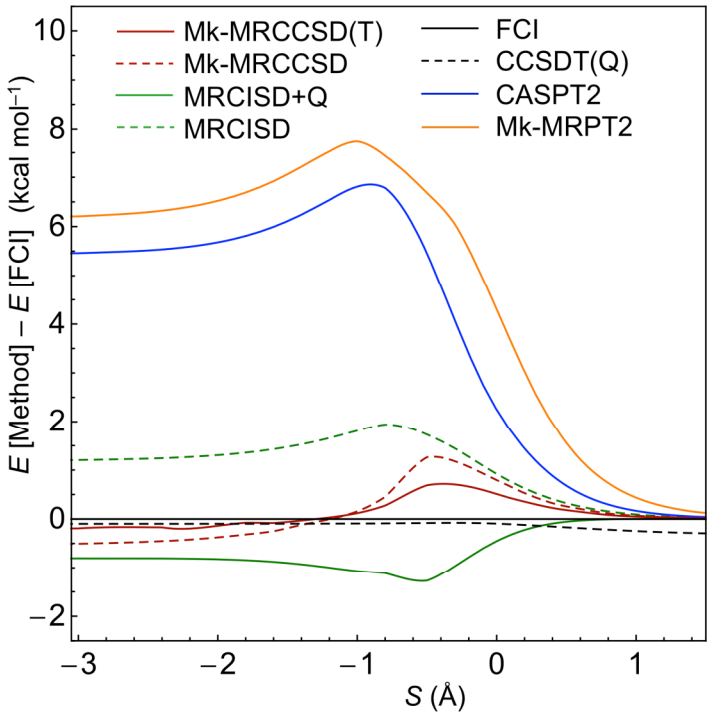


FIGURE 2.5: For the entrance channel of H + BeH, energy deviations are plotted with respect to FCI for different levels of theory with the cc-pVTZ basis set.

2.4.2 H + NH_2 → H_2 + NH

The H + NH_2 hydrogen abstraction reaction is a particularly interesting benchmark. The ground state of the NH_2 radical (C_{2v}) is $^2B_1$, as the unpaired electron occupies an out-of-plane b_1 orbital. When the H radical executes an in-plane attack on an N–H bond (Figure 2.1), the unpaired electron of NH_2 is thus perpendicular to the direction of abstraction. If the H radical approaches perpendicular to the plane of NH_2, recombination to NH_3 occurs rather than the radical-radical abstraction reaction of concern here. The in-plane attack yields a triplet ($^3A''$) surface for the H + NH_2 system that leads to ground-state H_2 + NH($^3\Sigma^-$) products and does not

require a multireference treatment. However, a corresponding open-shell singlet state (a^1A'') also occurs that correlates with H_2 + NH($^1\Delta$) and requires at least a two-reference treatment to properly describe the potential energy profile. Finally, if one starts with the lowest excited state of NH_2 having the unpaired electron in the molecular plane, then the reaction proceeds on a b^1A' surface. Our CASPT2/6-311G* computations show that at the $S = 0$ geometry (Figure 2.1), the a^1A'' and b^1A' states lie above the X^3A'' state by 0.42 and 1.46 eV, respectively; moreover, the $^1\Delta$ and $^1\Sigma^+$ states of the NH product radical are 1.73 eV and 2.82 eV, respectively, above the triplet ground state $^3\Sigma^-$ (vertical excitation energies).[94] Our benchmark multireference computations on H + NH_2 will thus focus on the a^1A'' surface.

The energy profiles at various levels of theory for hydrogen abstraction on the a^1A'' surface are shown in Figure 2.6, and corresponding errors with respect to FCI are plotted in Figure 2.7. The performance of each method generally follows the observations made in the H + BeH case. CASPT2 does well for the long-range interactions but substantially overestimates the hydrogen abstraction barrier in the intermediate region. Both multireference PT2 theories introduce an unphysical deflection in the energy curve near $S = -0.4$ Å that is not present in the MRCI and Mk-MRCC cases. The MRCISD curve exhibits a sizable 1–3 kcal mol^{-1} deviation from FCI, but this error is greatly reduced by the Davidson correction. MRCISD+Q and Mk-MRCCSD(T) both reproduce the FCI curve to better than 0.5 kcal mol^{-1}, but the latter method performs better for $S < 0.4$ Å.

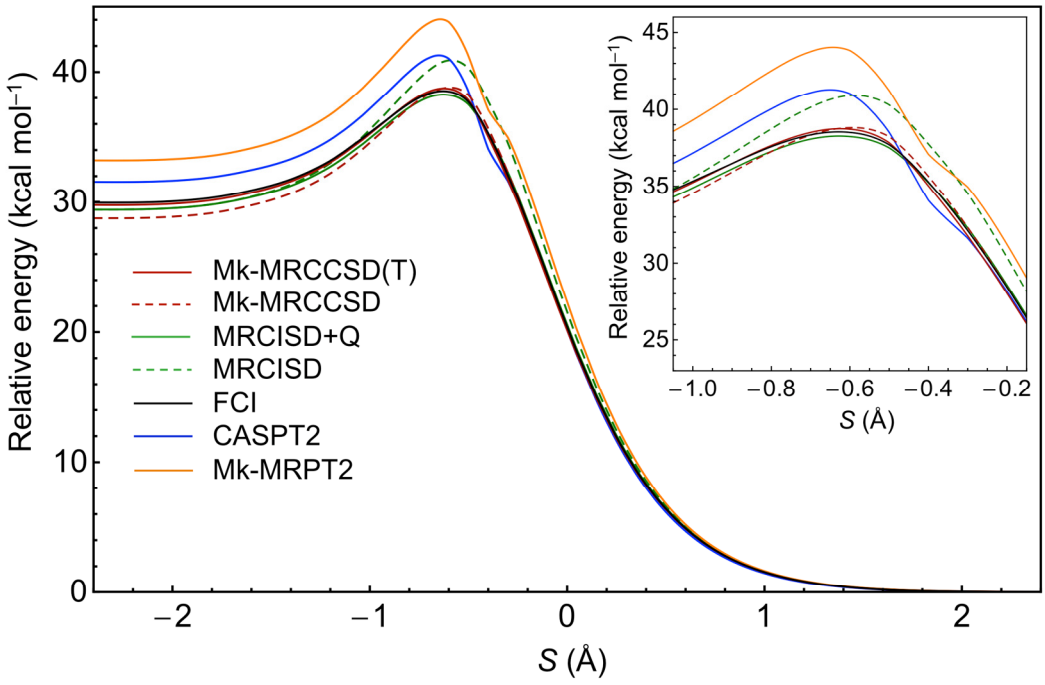


FIGURE 2.6: Potential energy profile on the a^1A'' surface for the H + NH_2 reaction computed at numerous levels of theory with the 6-311G* basis set. Separated [H + $NH_2(^2B_1)$, H_2 + $NH(^1\Delta)$] corresponds to $S = (+\infty,-\infty)$.

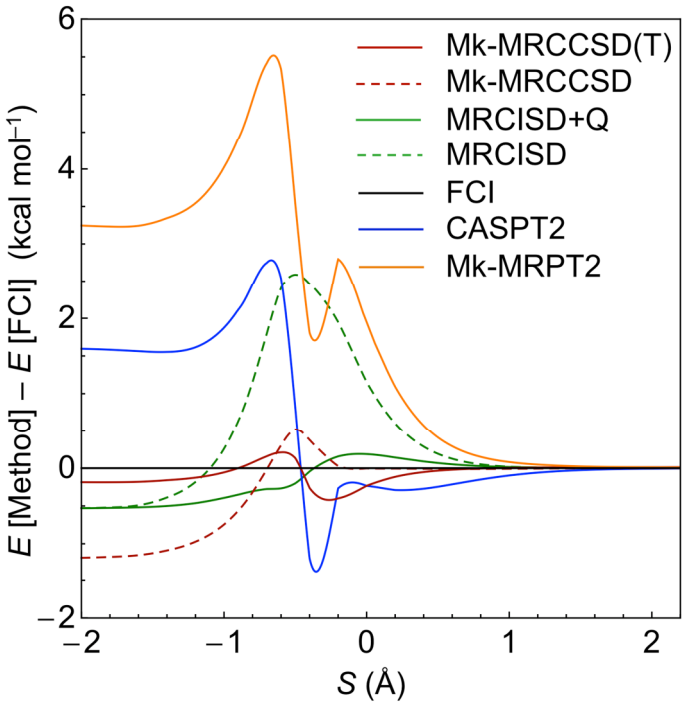


FIGURE 2.7: For the H + NH_2 reaction on the a^1A'' surface, energy deviations are plotted with respect to FCI for different levels of theory with the 6-311G* basis set.

2.4.3 $CH_3 + C_2H_5 \rightarrow CH_4 + C_2H_4$

The methyl + ethyl hydrogen abstraction reaction starts with two, separated doublet electronic states and produces closed-shell methane and ethylene products. Numerous results for the potential energy profile of the $CH_3 + C_2H_5$ abstraction reaction are shown in Figure 2.8. The overall reaction curve exhibits a similar shape as reaction 1 with the abstraction reaction energy –68.92 kcal mol^{-1}. The critical question is whether a barrier exists in the entrance channel. This is a particular concern because MRCISD+Q and CASPT2 give markedly different predictions. In the CBS limit, MRCISD gives a barrier of 1.2 kcal mol^{-1} with respect to the pre-reaction

complex. With inclusion of the +Q correction, the MRCISD+Q barrier is only about 0.1 kcal mol^{-1} above the complex minimum. In contrast, no barrier exists in the CASPT2 curves, which display the deepest interaction potentials for each value of *S*.

As expected, single-reference CCSD(T) theory produces unphysical results as the system evolves in the reverse direction from $CH_4 + C_2H_4$ to separated radicals. The most rigorous Mk-MRCCSD(T) method yields a curve that monotonically decreases as the reactants approach, with no barrier in the entrance channel at all. This key conclusion bodes well for the importance of radical-radical abstractions in combustion chemistry. Figure 2.9 shows the deviation of the binding energy at the various levels of theory with respect to Mk-MRCCSD(T), which gave the most accurate results compared to FCI for the H + BeH and $H + NH_2$ benchmarks. The CASPT2 predictions appear within 0.5 kcal mol^{-1} of the Mk-MRCCSD(T) results, but overestimates the binding energy. The MRCISD+Q curve is quite poor for this reaction, deviating in the positive direction by more than 3 kcal mol^{-1} from the benchmark. Although not shown in the figures, the Mk-MRCCSD(T) and single-reference CCSD(T) curves are virtually identical when $S < -0.1$ Å. Therefore for the CBS extrapolated energies in reactions 3, 4 and 5, we only report the Mk-MRCCSD(T) energies in the multireference region.

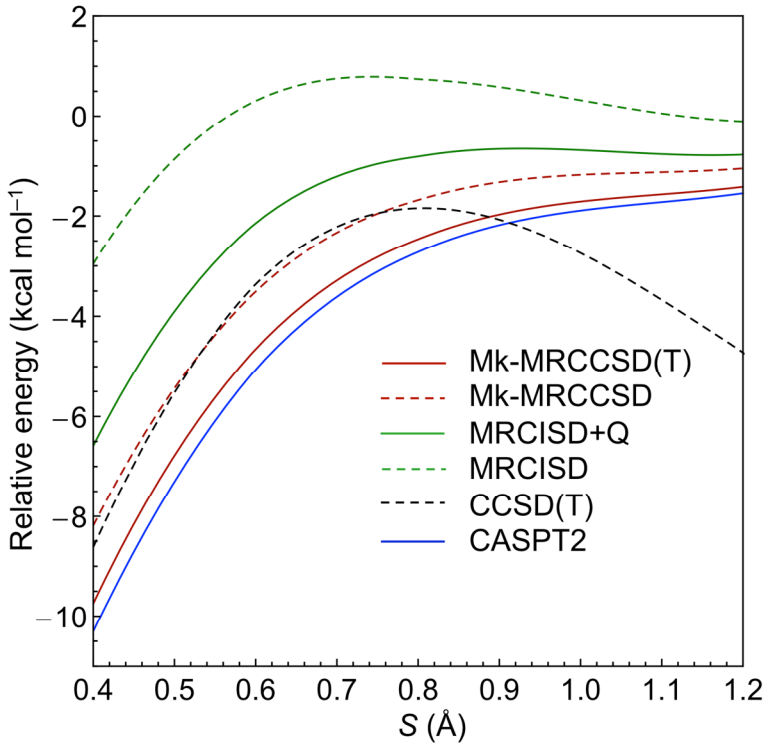


FIGURE 2.8: Potential energy curves for the entrance channel of $CH_3 + C_2H_5$ computed at numerous levels of theory and extrapolated to the CBS limit.

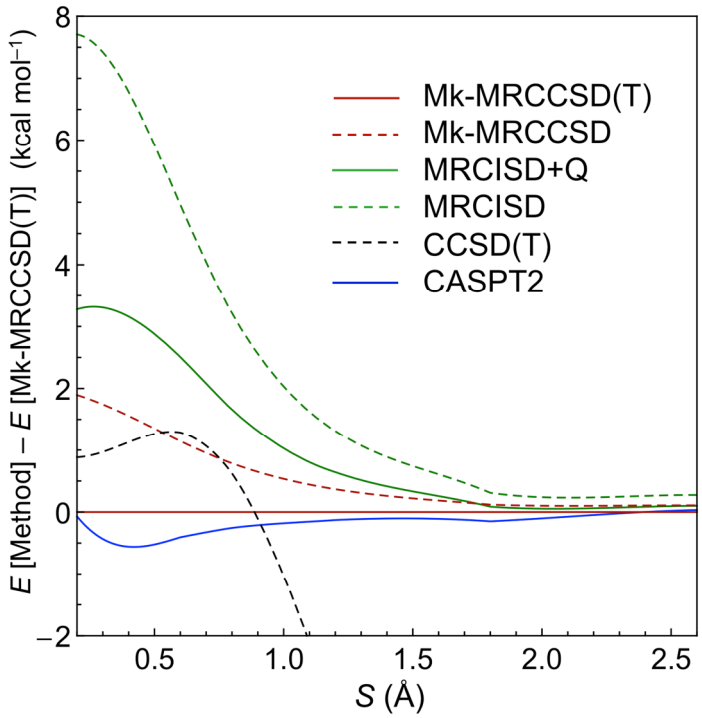


FIGURE 2.9: For the entrance channel of $CH_3 + C_2H_5$, energy deviations are plotted with respect to the Mk-MRCCSD(T) benchmark; CBS extrapolations were employed for all methods.

Table 3. Focal point analysis[a] for the $CH_3 + C_2H_5$ reaction at $S = 0.4$ Å

	ΔE_e[CAS(2,2)]	$+\delta$ [Mk-MRCCSD]	$+\delta$ [Mk-MRCCSD(T)]	ΔE_{NET}
cc-pVTZ	+8.16	–15.72	–1.33	–8.89
cc-pVQZ	+8.25	–16.17	–1.46	–9.39
cc-pV5Z	+8.29	[–16.33]	[–1.51]	[–9.55]
CBS	[+8.31]	[–16.50]	[–1.56]	[–9.75]

[a] δ denotes the *increment* in the energy difference (ΔE_e) with respect to the previous level of theory in the hierarchy CAS(2,2) → Mk-MRCCSD → Mk-MRCCSD(T). (Unbracketed, bracketed) numbers result from (explicit computations, basis set extrapolations with eqs 12 and 13). The focal-point table targets ΔE_e[Mk-MRCCSD(T)] in the complete basis set limit (NET/CBS).

2.4.4 $H + C_2H_5 \rightarrow H_2 + C_2H_4$

The H + ethyl reaction is a simpler example of the same electronic structure transformations encountered for methyl + ethyl, and similarly begins as two, separated doublet electronics states and produces closed-shell products. In contrast to methyl + ethyl, all the levels of theory predict no barrier for H + ethyl hydrogen abstraction at the CBS limit, as shown in Figure 2.10. The potential energy curve monotonically decreases with no long-range van der Waals complex and overall –68.54 kcal mol^{-1} energy of reaction. The benchmark Mk-MRCCSD(T) curve displays the greatest binding, and CASPT2 closely mimics this result. In the range S = 0.0–1.2 Å, the binding energy difference between CASPT2 and Mk-MRCCSD(T) is less than 0.5 kcal mol^{-1} (Figure 2.11). In this same range, MRCISD performs poorly, predicting binding energies that are too small and in error by almost 5 kcal mol^{-1} at S = 0. While the Davidson correction once again improves MRCISD considerably, MRCISD+Q still fails to give quantitatively reliable predictions. The barrierless and steeply descending entrance channel given by Mk-MRCCSD(T)/CBS theory enhances the argument for the importance of radical-radical abstraction reactions in combustion.

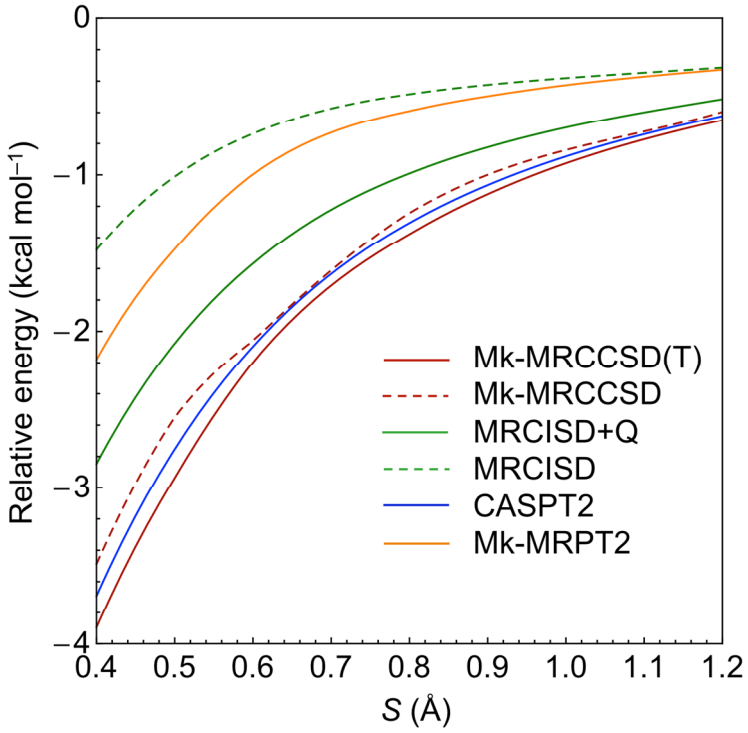


FIGURE 2.10: Potential energy curves for the entrance channel of H + C_2H_5 computed at numerous levels of theory and extrapolated to the CBS limit.

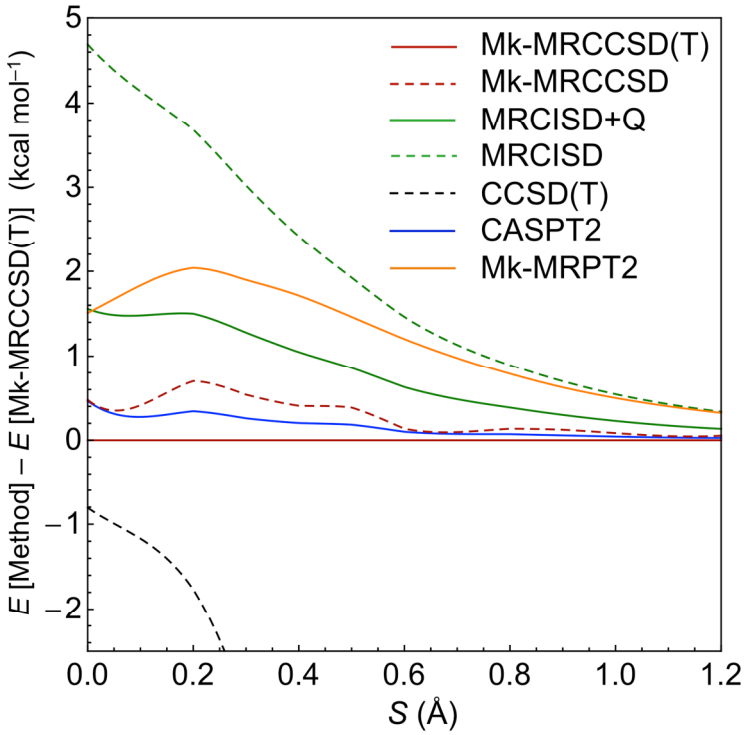


FIGURE 2.11: For the entrance channel of H + C_2H_5, energy deviations are plotted with respect to the Mk-MRCCSD(T) benchmark; CBS extrapolations were employed for all methods.

TABLE 2.4: Focal point analysis[a] for the H + C_2H_5 reaction at S = 0.4 Å

	ΔE_e[CAS(2,2)]	$+\delta$[Mk-MRPT2]	$+\delta$[Mk-MRCCSD]	$+\delta$[Mk-MRCCSD(T)]	ΔE_{NET}
cc-pVTZ	+3.10	–4.85	–1.25	–0.33	–3.32
cc-pVQZ	+3.10	–5.10	–1.28	–0.37	–3.65
cc-pV5Z	+3.10	[–5.19]	[–1.29]	[–0.39]	[–3.77]
CBS	[+3.10]	[–5.28]	[–1.31]	[–0.41]	[–3.89]

[a] For notation, see footnote *a* of Table 3.

[b] The core correlation correction for the binding energy of this geometry is 0.02 kcal mol^{-1}. In the consideration of computational expense, only valence electrons were correlated in the reaction energy profile.

2.4.5 $H + HCO \rightarrow H_2 + CO$

The hydrogen abstraction reaction between formyl radical and hydrogen radical is depicted in Figure 2.1. All the multireference methods, regardless of basis set, predict that no abstraction barrier exists the entrance channel of H + HCO. An initial series of computations (Figure 2.12) was performed with the cc-pVDZ(*sp*) basis set. This pruned basis set allowed FCI computations to be performed for H + HCO. The final energy profiles (Figure 2.13) were obtained via CBS extrapolations at each level of theory. The reaction energy is predicted to be –88.63 kcal mol^{-1} at CCSD(T)/CBS using our extrapolation approach.

In Figure 2.12, CASPT2 is almost indistinguishable from FCI for $S > 0.2$ Å, while Mk-MRCCSD(T) yields a curve shifted slightly to higher energy. However, when the basis set is increased from cc-pVDZ(*sp*) to the CBS limit, CASPT2 exhibits an underestimation of the binding energy. The CASPT2/cc-pVDZ(*sp*) curve (blue) shifts from the negative side of Mk-MRCCSD(T) (red) in Figure 2.12 to the positive side of both Mk-MRCCSD(T) and MRCISD+Q (green) in Figure 2.13. Similar shifting behavior is also observed for the methyl + ethyl and H + ethyl reactions when the basis set increases from cc-pVDZ to cc-pVQZ. Note that the single-reference CCSDT(Q) curve in Figure 2.12 maintains accuracy much longer than CCSD(T) as *S* increases, but the perturbative nature of the (Q) treatment eventually causes errors greater than 1 kcal mol^{-1} past 1.0 Å.

The energy deviation plots (Figure 2.14, cc-pVDZ(*sp*); Figure 2.15, CBS) demonstrate that CASPT2 perturbation theory generates excellent results for the long-range H + HCO interaction but deteriorates markedly at shorter range ($S < 0.2$ Å). In contrast, the error of Mk-MRCCSD(T)/cc-pVDZ(*sp*) with respect to FCI is less than 0.6 kcal mol^{-1} over the whole range

of the plot in Figure 2.14 ($S > -0.3$ Å). In this comparison Mk-MRCCSD(T) clearly outperforms both MRCISD and MRCISD+Q. In Figure 2.15 the deviation between MRCISD+Q and Mk-MRCCSD(T) does not exceed 0.5 kcal mol^{-1}, but without the Davidson correction the disparity of MRCISD can be larger than 2 kcal mol^{-1}.

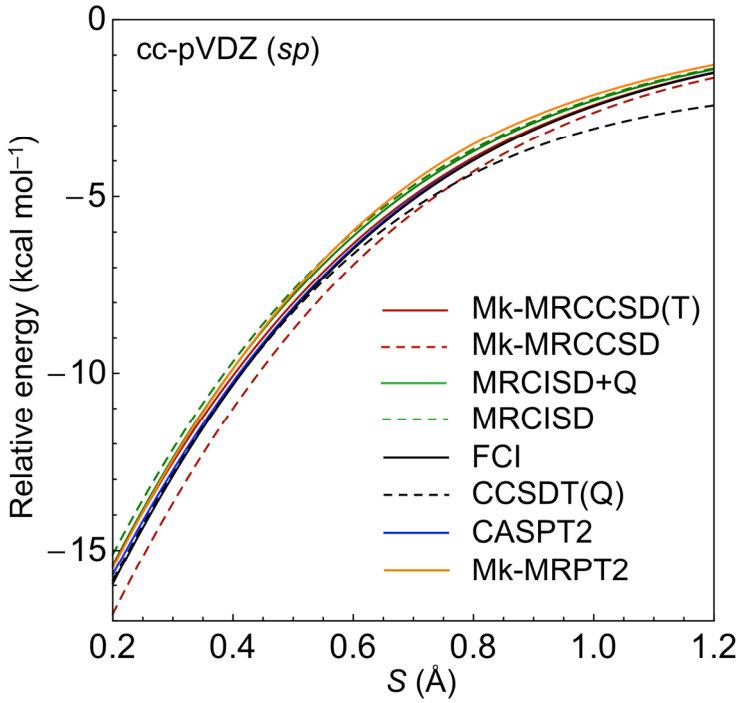


FIGURE 2.12: Potential energy curves for the entrance channel of H + HCO computed at numerous levels of theory using a cc-pVDZ(*sp*) basis set.

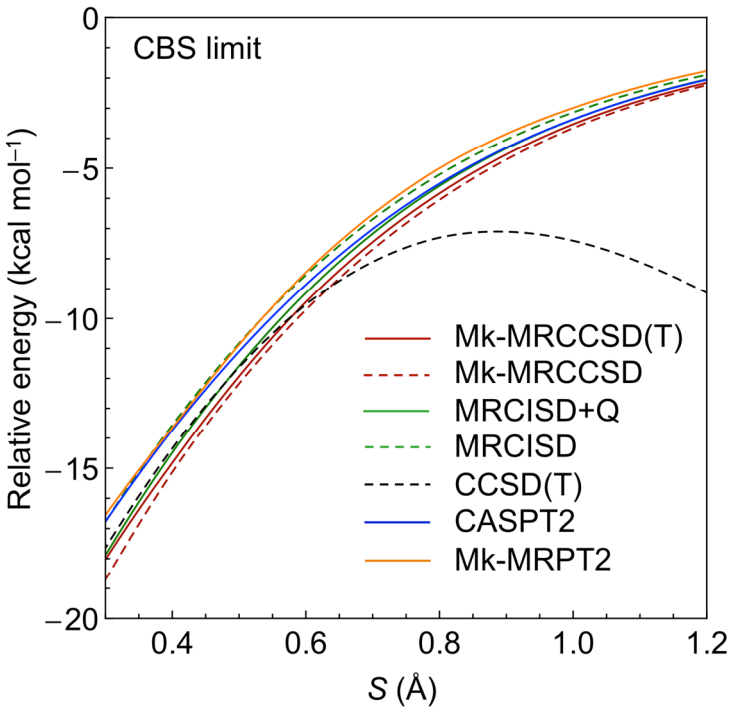


FIGURE 2.13: Potential energy curves for the entrance channel of H + HCO computed at numerous levels of theory and extrapolated to the CBS limit.

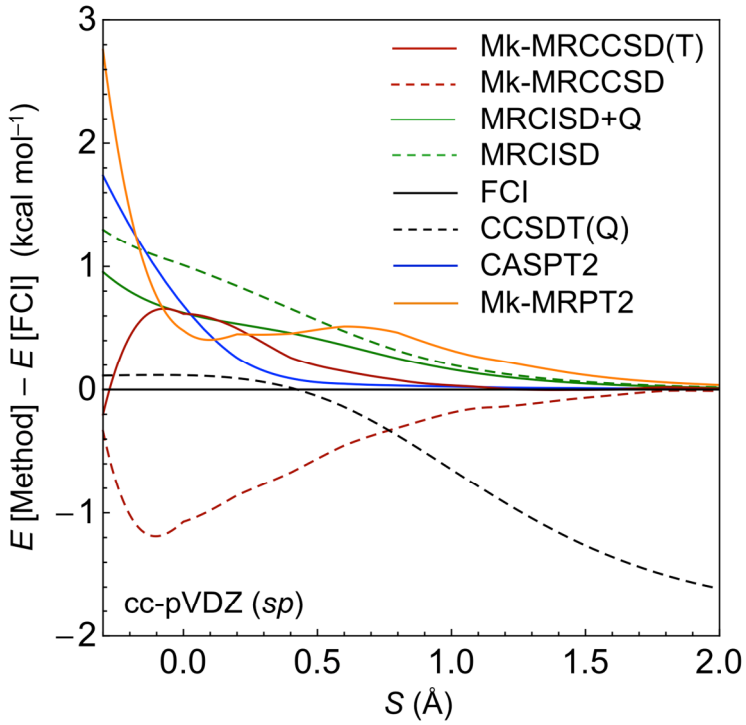


FIGURE 2.14: For the entrance channel of H + HCO, energy deviations are plotted with respect to FCI for different levels of theory with the cc-pVDZ(*sp*) basis set.

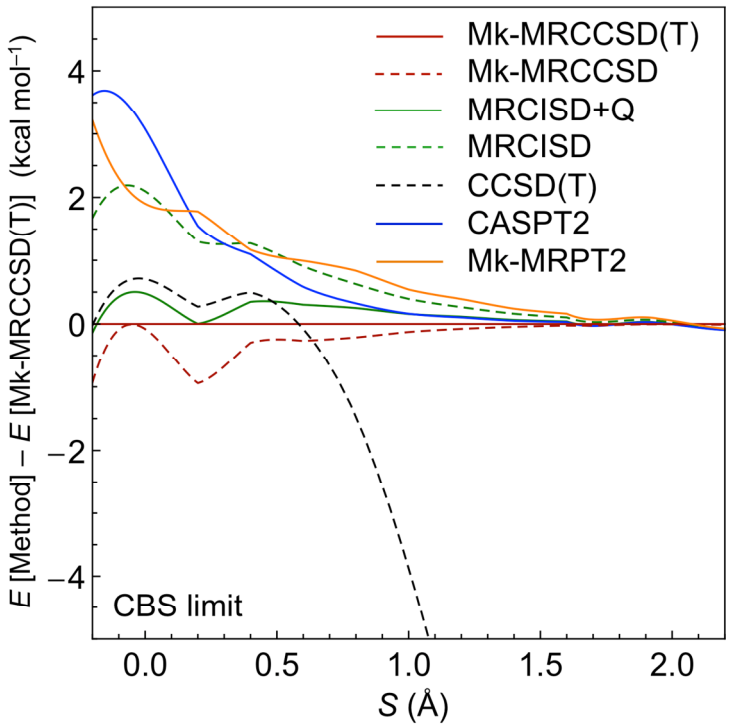


FIGURE 2.15: For the entrance channel of H + HCO, energy deviations are plotted with respect to the Mk-MRCCSD(T) benchmark; CBS extrapolations were employed for all methods.

TABLE 2.5: Focal point analysis[a] for the H + HCO reaction at S = 0.4 Å

	ΔE_e[CAS(2,2)]	$+\delta$[Mk-MRPT2]	$+\delta$[Mk-MRCCSD]	$+\delta$[Mk-MRCCSD(T)]	ΔE_{NET}
cc-pVTZ	–9.98	–3.05	–2.09	+0.61	–14.51
cc-pVQZ	–9.92	–3.51	–1.73	+0.43	–14.74
cc-pV5Z	–9.85	[–3.68]	[–1.60]	[+0.37]	[–14.77]
CBS	[–9.81]	[–3.85]	[–1.46]	[+0.30]	[–14.83]

[a] For notation, see footnote *a* of Table 3.

[b] The core correlation correction for the binding energy of this geometry is 0.013 kcal mol^{-1}. In the consideration of computational expense, only valence electrons were correlated in the reaction energy profile.

2.5 CONCLUSIONS

Our study demonstrates the accuracy of Mk-MRCC theory. With the use of localized orbitals, the Mk-MRCCSD(T) energies deviate from FCI by less than 1 kcal mol^{-1} in reactions 1, 2 and 5. Tikhonov regularization with extrapolation to $\omega = 0$ was successful in overcoming Mk-MRCC convergence difficulties encountered at some geometries. In the middle-to-long range region of the potential energy surface, our computations found the performance of CASPT2 is very good in comparison to Mk-MRCCSD(T). However, in the region where electron transfer takes place, the results from perturbation theory (Mk-MRPT2 and CASPT2) are less reliable. We have reported the rigorous one-dimensional energy profile with the Mk-MRCCSD(T) energy extrapolated to the CBS limit and showed that the hydrogen abstractions for H + HCO, H + C_2H_5, and CH_3 + C_2H_5 are all barrierless processes. The ZPVE corrections along the reaction path are also provided.

2.6 REFERENCES

(1) Howard, J. B. *Symposium (International) on Combustion* **1991**, *23*, 1107-1127.

(2) Benson, S. W. *Can. J. Chem.* **1983**, *61*, 881-887.

(3) Klippenstein, S. J.; Harding, L. B. *J. Phys. Chem. A* **1999**, *103*, 9388-9398.

(4) Klippenstein, S. J.; Georgievskii, Y.; Harding, L. B. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1133-1147.

(5) Harding, L. B.; Klippenstein, S. J.; Jasper, A. W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4055-4070.

(6) Gibian, M. J.; Corley, R. C. *Chem. Rev.* **1973**, *73*, 441-464.

(7) Benson, S. W. *J. Phys. Chem.* **1985**, *89*, 4366-4369.

(8) Fahr, A.; Laufer, A.; Klein, R.; Braun, W. *J. Phys. Chem.* **1991**, *95*, 3218-3224.

(9) Fahr, A.; Laufer, A. H. *J. Phys. Chem.* **1990**, *94*, 726-729.

(10) Fahr, A.; Laufer, A. H. *J. Phys. Chem.* **1995**, *99*, 262-264.

(11) Altinay, G.; Macdonald, R. G. *J. Phys. Chem. A* **2013**, *118*, 38-54.

(12) Bahng, M.-K.; Macdonald, R. G. *J. Phys. Chem. A* **2007**, *111*, 3850-3861.

(13) Bedjanian, Y.; Le Bras, G.; Poulet, G. *J. Phys. Chem. A* **1999**, *103*, 7017-7025.

(14) Sangwan, M.; Chesnokov, E. N.; Krasnoperov, L. N. *J. Phys. Chem. A* **2012**, *116*, 6282-6294.

(15) Sangwan, M.; Krasnoperov, L. N. *J. Phys. Chem. A* **2012**, *116*, 11817-11822.

(16) Braunstein, M.; Panfili, R.; Shroll, R.; Bernstein, L. *J. Chem. Phys.* **2005**, *122*, 184307.

(17) Deyerl, H.-J.; Clements, T. G.; Luong, A. K.; Continetti, R. E. *J. Chem. Phys.* **2001**, *115*, 6931-6940.

(18) Harding, L. B. *J. Phys. Chem.* **1991**, *95*, 8653-8660.

(19) Harding, L. B.; Wagner, A. F. *Symposium (International) on Combustion* **1989**, *22*, 983-989.

(20) Karkach, S. P.; Osherov, V. I. *J. Chem. Phys.* **1999**, *110*, 11918-11927.

(21) Maergoiz, A. I.; Nikitin, E. E.; Troe, J. *J. Chem. Phys.* **1995**, *103*, 2083-2091.

(22) Tranter, R. S.; Klippenstein, S. J.; Harding, L. B.; Giri, B. R.; Yang, X.; Kiefer, J. H. *J. Phys. Chem. A* **2010**, *114*, 8240-8261.

(23) Harding, L. B.; Klippenstein, S. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3016-3020.

(24) Heazlewood, B. R.; Jordan, M. J.; Kable, S. H.; Selby, T. M.; Osborn, D. L.; Shepler, B. C.; Braams, B. J.; Bowman, J. M. *Proc. Natl. Acad. Sci.* **2008**, *105*, 12719-12724.

(25) Houston, P.; Kable, S. *Proc. Natl. Acad. Sci.* **2006**, *103*, 16079-16082.

(26) Sivaramakrishnan, R.; Michael, J. V.; Wagner, A. F.; Dawes, R.; Jasper, A. W.; Harding, L. B.; Georgievskii, Y.; Klippenstein, S. J. *Combust. Flame* **2011**, *158*, 618-632.

(27) Sivaramakrishnan, R.; Michael, J. V.; Harding, L. B.; Klippenstein, S. J. *J. Phys. Chem. A* **2012**, *116*, 5981-5989.

(28) Sivaramakrishnan, R.; Su, M. C.; Michael, J. V.; Klippenstein, S. J.; Harding, L. B.; Ruscic, B. *J. Phys. Chem. A* **2011**, *115*, 3366-3379.

(29) Townsend, D.; Lahankar, S.; Lee, S.; Chambreau, S.; Suits, A.; Zhang, X.; Rheinecker, J.; Harding, L.; Bowman, J. *Science* **2004**, *306*, 1158-1161.

(30) Dürrstein, S. H.; Olzmann, M.; Aguilera-Iparraguirre, J.; Barthel, R.; Klopper, W. *Chem. Phys. Lett.* **2011**, *513*, 20-26.

(31) Magoon, G. R.; Aguilera-Iparraguirre, J.; Green, W. H.; Lutz, J. J.; Piecuch, P.; Wong, H. W.; Oluwole, O. O. *Int. J. Chem. Kinet.* **2012**, *44*, 179-193.

(32) Zhu, R.; Xu, Z.; Lin, M. *J. Chem. Phys.* **2004**, *120*, 6566-6573.

(33) Mousavipour, S. H.; Homayoon, Z. *J. Phys. Chem. A* **2003**, *107*, 8566-8574.

(34) Klippenstein, S. J.; Georgievskii, Y.; Harding, L. B. *J. Phys. Chem. A* **2011**, *115*, 14370-14381.

(35) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *Mol. Phys.* **1998**, *94*, 157-171.

(36) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *J. Chem. Phys.* **1999**, *110*, 6171-6188.

(37) Evangelista, F. A.; Allen, W. D.; Schaefer III, H. F. *J. Chem. Phys.* **2006**, *125*, 154113.

(38) Evangelista, F. A.; Allen, W. D.; Schaefer III, H. F. *J. Chem. Phys.* **2007**, *127*, 024102.

(39) Evangelista, F. A.; Prochnow, E.; Gauss, J.; Schaefer III, H. F. *J. Chem. Phys.* **2010**, *132*, 074107.

(40) Evangelista, F. A.; Simmonett, A. C.; Allen, W. D.; Schaefer III, H. F.; Gauss, J. *J. Chem. Phys.* **2008**, *128*, 124104.

(41) Evangelista, F. A.; Simmonett, A. C.; Schaefer III, H. F.; Mukherjee, D.; Allen, W. D. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4728-4741.

(42) Prochnow, E.; Evangelista, F. A.; Schaefer III, H. F.; Allen, W. D.; Gauss, J. *J. Chem. Phys.* **2009**, *131*, 064109.

(43) Jeziorski, B.; Monkhorst, H. J. *Phys. Rev. A* **1981**, *24*, 1668.

(44) Köhn, A.; Hanauer, M.; Mueck, L. A.; Jagau, T. C.; Gauss, J. *WIREs Comput. Mol. Sci.* **2013**, *3*, 176-197.

(45) Wu, J. I.-C.; Mo, Y.; Evangelista, F. A.; Schleyer, P. v. R. *Chem. Comm.* **2012**, *48*, 8437-8439.

(46) Lu, T.; Hao, Q.; Simmonett, A. C.; Evangelista, F. A.; Yamaguchi, Y.; Fang, D.-C.; Schaefer III, H. F. *J. Phys. Chem. A* **2010**, *114*, 10850-10856.

(47) Lu, T.; Simmonett, A. C.; Evangelista, F. A.; Yamaguchi, Y.; Schaefer III, H. F. *J. Phys. Chem. A* **2009**, *113*, 13227-13236.

(48) Saito, T.; Nishihara, S.; Kataoka, Y.; Nakanishi, Y.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Okumura, M.; Yamaguchi, K. *J. Phys. Chem. A* **2010**, *114*, 12116-12123.

(49) Saito, T.; Yasuda, N.; Nishihara, S.; Yamanaka, S.; Kitagawa, Y.; Kawakami, T.; Okumura, M.; Yamaguchi, K. *Chem. Phys. Lett.* **2011**, *505*, 11-15.

(50) Schreiner, P. R.; Reisenauer, H. P.; Ley, D.; Gerbig, D.; Wu, C.-H.; Allen, W. D. *Science* **2011**, *332*, 1300-1303.

(51) Schreiner, P. R.; Reisenauer, H. P.; Pickard Iv, F. C.; Simmonett, A. C.; Allen, W. D.; Mátyus, E.; Császár, A. G. *Nature* **2008**, *453*, 906-909.

(52) Šimsa, D.; Demel, O.; Bhaskaran-Nair, K.; Hubač, I.; Mach, P.; Pittner, J. *Chem. Phys.* **2012**, *401*, 203-207.

(53) Andersson, K.; Malmqvist, P. Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218-1226.

(54) Dyall, K. G. *J. Chem. Phys.* **1995**, *102*, 4909-4918.

(55) Knowles, P. J.; Werner, H.-J. *Chem. Phys. Lett.* **1985**, *115*, 259-267.

(56) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1985**, *82*, 5053-5063.

(57) Knowles, P. J.; Werner, H.-J. *Chem. Phys. Lett.* **1988**, *145*, 514-522.

(58) Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*, 5803-5814.

(59) Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, *8*, 61-72.

(60) Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2000**, *112*, 5546-5557.

(61) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256-4266.

(62) Crawford, T. D.; Schaefer III, H. F. *Rev. Comput. Chem.* **2000**, *14*, 33-136.

(63) *MRCC, a quantum chemical program suite written by M. Kállay, Z. Rolik, I. Ladjánszki, L. Szegedy, B. Ladóczki, J. Csontos, and B. Kornis. See also Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, and M. Kállay J. Chem. Phys. 139 094105 (**2013**) as well as: www.mrcc.hu.*

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

(65) Knowles, P. J.; Handy, N. C. *Chem. Phys. Lett.* **1984**, *111*, 315-321.

(66) Knowles, P. J.; Handy, N. C. *Comput. Phys. Commun.* **1989**, *54*, 75-83.

(67) Sherrill, C. D.; Leininger, M. L.; Van Huis, T. J.; Schaefer III, H. F. *J. Chem. Phys.* **1998**, *108*, 1040-1049.

(68) Sherrill, C. D.; Van Huis, T. J.; Yamaguchi, Y.; Schaefer III, H. F. *J. Mol. Struct.-THEOCHEM* **1997**, *400*, 139-156.

(69) Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007-1023.

(70) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650-654.

(71) McLean, A.; Chandler, G. *J. Chem. Phys.* **2008**, *72*, 5639-5648.

(72) Allen, W. D.; East, A. L. L.; Császár, A. G. In *Structure and Conformations of Non-rigid Molecules*; Laane, J., Dakkouri, M., van der Veken, B. and Oberhammer, H., Eds.; Kluwer: Dordrecht: The Netherlands, 1993, p 343-373.

(73) Császár, A. G.; Tarczay, G.; Leininger, M. L.; Polyansky, O. L.; Tennyson, J.; Allen, W. D. In *Spectroscopy from Space*; J. Demaison and Sarka, K., Eds.; Kluwer: Dordrecht: The Netherlands, 2001, p 317-339.

(74) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638-4650.

(75) Gonzales, J. M.; Allen, W. D.; Schaefer III, H. F. *J. Phys. Chem. A* **2005**, *109*, 10613-10628.

(76) Gonzales, J. M.; Pak, C.; Cox, R. S.; Allen, W. D.; Schaefer III, H. F.; Császár, A. G.; Tarczay, G. *Chem.-Eur. J.* **2003**, *9*, 2173-2192.

(77) *For reaction 9, 10 and 11, the core correlation correction in Mk-MRCCSD(T)/cc-pVTZ level of theory for binding energy at the geometry S = 0.4 Angstrom is less than XX kcal* mol^{-1}.

(78) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104-6114.

(79) Feller, D. *J. Chem. Phys.* **1993**, *98*, 7059-7071.

(80) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639-9646.

(81) *Evangelista F. A. and Simmonett A. C. , PSIMRCC, a computer code written at the University of Georgia and capable of performing MRCC computations, see http://sirius.chem.vt.edu/psi4manual/4.0b5/psimrcc.html (**2007**).*

(82) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P. *J. Comput. Chem.* **2007**, *28*, 1610-1616.

(83) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556-565.

(84) *MOLPRO is a package of ab initio programs written by H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang .*

(85) Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215-223.

(86) Allen, W. D.; Császár, A. G. *J. Chem. Phys.* **1993**, *98*, 2983-3015.

(87) Allen, W. D.; Bodi, A.; Szalay, V.; Császár, A. G. *J. Chem. Phys.* **2006**, *124*, 224310.

(88) Das, S.; Kállay, M.; Mukherjee, D. *Chem. Phys.* **2012**, *392*, 83-89.

(89) Boys, S. *Proc. R. Soc. Lond., A* **1960**, *258*, 402-411.

(90) Das, S.; Mukherjee, D.; Kállay, M. *J. Chem. Phys.* **2010**, *132*, 074103.

(91) *Evangelista, F.A. Personal communication. (**2011**).*

(92) *The differences between the extrapolation energies and the energies computed for* $\omega = 0$ *are placed in the supporting information.*

(93) Purvis, G. D.; Shepard, R.; Brown, F. B.; Bartlett, R. J. *Int. J. Quantum Chem.* **1983**, *23*, 835-845.

(94) Owono, L. O.; Jaidane, N.; Njock, M. K.; Lakhdar, Z. B. *J. Chem. Phys.* **2007**, *126*, 244302.

CHAPTER 3

ANOMALOUS IMAGINARY FREQUENCIES IN CYCLIC AND LINEAR SYSTEMS DUE TO INSUFFICIENT BASIS SETS [†]

[†] D. Brandon Magers, Zachary Schneiderman, and Wesley D. Allen. To be submitted to *Journal of Chemical Theory and Computation*.

3.1 ABSTRACT

Imaginary or low harmonic vibrational frequencies are obtained for a number of planar, linear, and quasilinear systems when electron correlation methods are applied with an array of Pople basis sets. These anomalous results are due to insufficiencies in the basis set and not caused by the level of theory, as we report divergent results even at the CCSD(T) level. The disparities dissipate with the use of correlation-consistent basis sets such as the Dunning or Atomic Natural Orbital series.

3.2 INTRODUCTION

In 2006, Moran *et al.*[1] found non-planar minima for benzene and other arenes when performing standard quantum chemical computations using Pople basis sets. Before it was known that insufficient basis sets produced non-planar minima, it was first established by several research groups that these basis sets produced erroneously low harmonic vibrational frequency values. Benzene was found to have over a 30% error in the out of plane ω_4 (b_{2g}) bending frequency at the MP2/6-311G(d,p) level of theory when compared to experimental results.[2] Even at the higher CCSD(T) level of theory, two anomalously low out-of-plane bending frequencies were found for $\omega_4(b_{2g})$ and $\omega_5(b_{2g})$.[3] The ANO basis set series was highly recommended to correct for this problem. Results such as these for benzene have been documented for phenol, benzaldehyde, and salicylaldehyde.[4] At the MP2/6-311++G(d,p) level of theory, a computed (b_1) frequency of 228 cm^{-1} for pyridine is much lower than the experimental value of 745 cm^{-1}.[5] The results by Moran *et al.* followed the same trend with a variety of arenes including benzene,

but also found something more alarming altogether, reporting the first computed imaginary frequencies for benzene, suggesting a non-planar minimum despite its long established planarity in both experimental and computational works. In recent years, similar non-planar frequencies have been found for the nitrogenous bases in DNA and RNA.[6] Even for the less common arene $(B_6C)^{2-}$ anion, this anomalous error appears.[7] In each case the non-planar nature of the molecule was typified by the presence of imaginary, out-of-plane bending frequencies.[1,6,7] This is not a problem when using single determinant methods such as Hartree-Fock (HF); however, these imaginary frequencies appear when using post-HF correlated methods such as Møller-Plesset second-order perturbation theory (MP2), configuration interaction theory (CI), and coupled cluster theory (CC).

Several solutions to this problem have been proposed. The first proposal is the use of different basis sets altogether.[1] The use of different, higher-level basis sets can reduce the basis set incompleteness error (BSIE) and the basis set superposition error (BSSE), which truly stand at the root of the problem. Saturating the basis sets with *s* and *p* valence orbitals and *d* polarization functions is not enough; these functions must all be included in balance and preferably with the additional inclusion of higher angular momentum functions such as *f* orbitals.[1] Possible choices include the correlation consistent Dunning[8] or atomic natural orbital (ANO) basis sets[9] which are both optimized to minimize the BSIE with post-Hartree-Fock methods and did correct the problem for benzene and several other arenes.[1,3,10] It was expected that this trend would continue for the molecules investigated in this paper. In order to verify this prediction, the use of these basis sets was probed for each molecule in this paper. A second proposed solution is the inclusion of a specific BSSE correction.[11] It is proposed that the non-planarity problem is truly rooted in intramolecular BSSE. Therefore a typical correction for such

an error, the counterpoise (CP) correction, should fix the problem. Such corrections are typically used for intermolecular interactions where BSSEs are more often seen; however, a typical CP correction led to the elimination of imaginary frequencies and generally produced results lining up much closer with experimental data. This work chose not to take this approach due to the difficulty of precisely designating fragments within a single molecule, and thus this method is not very systematic.

This paper examines these anomalous results at higher levels of theory, as well as for linear and quasilinear molecules. To this end the harmonic frequencies of several linear, quasilinear, and planar molecules were examined at MP2, CCSD, and CCSD(T) levels of theory for a large variety of Pople basis sets. Computations were also run for Dunning and atomic natural orbital basis sets in order to see if this provided a solution to the problem as would be predicted. To this end we illuminate which basis sets can avoid these anomalous results while managing the computational cost. We hope to ensure the greatest accuracy and efficiency of future computational works on related molecules.

3.3 THEORETICAL METHODS AND COMPUTATIONAL DETAILS

The molecules displayed in Figure 3.1 were characterized using Møller-Plesset second-order perturbation theory (MP2), coupled cluster theory with single and double excitations (CCSD), and coupled cluster theory with single excitations, double excitations, and perturbative treatment of triple excitations (CCSD(T)). All computations did not correlate the core electrons. Acetylene, diacetylene, cyanoacetylene, cyanogen, and triacetylene were constrained to a linear geometry; butatriene and ethenone (ketene) were constrained to linear heavy-atom frameworks;

and benzene, furan, thiophene, aniline, phenol, and naphthalene were constrained to have planar rings. Each of these constraints represented the energy minimum as previously found by experiment. A geometry optimization and harmonic frequency computation were then executed for each of these molecules at each level of theory using the basis sets listed in Table 3.1. The basis sets utilized include multiple Pople,[12] Dunning,[8,13] and ANO[9] basis sets. All computations were run with the CFOUR program package.[14] Since the geometry was constrained, any imaginary frequencies found during these computations demonstrate that the planar or linear structure is not a true minimum, despite a preponderance of theoretical and experimental evidence to the contrary.

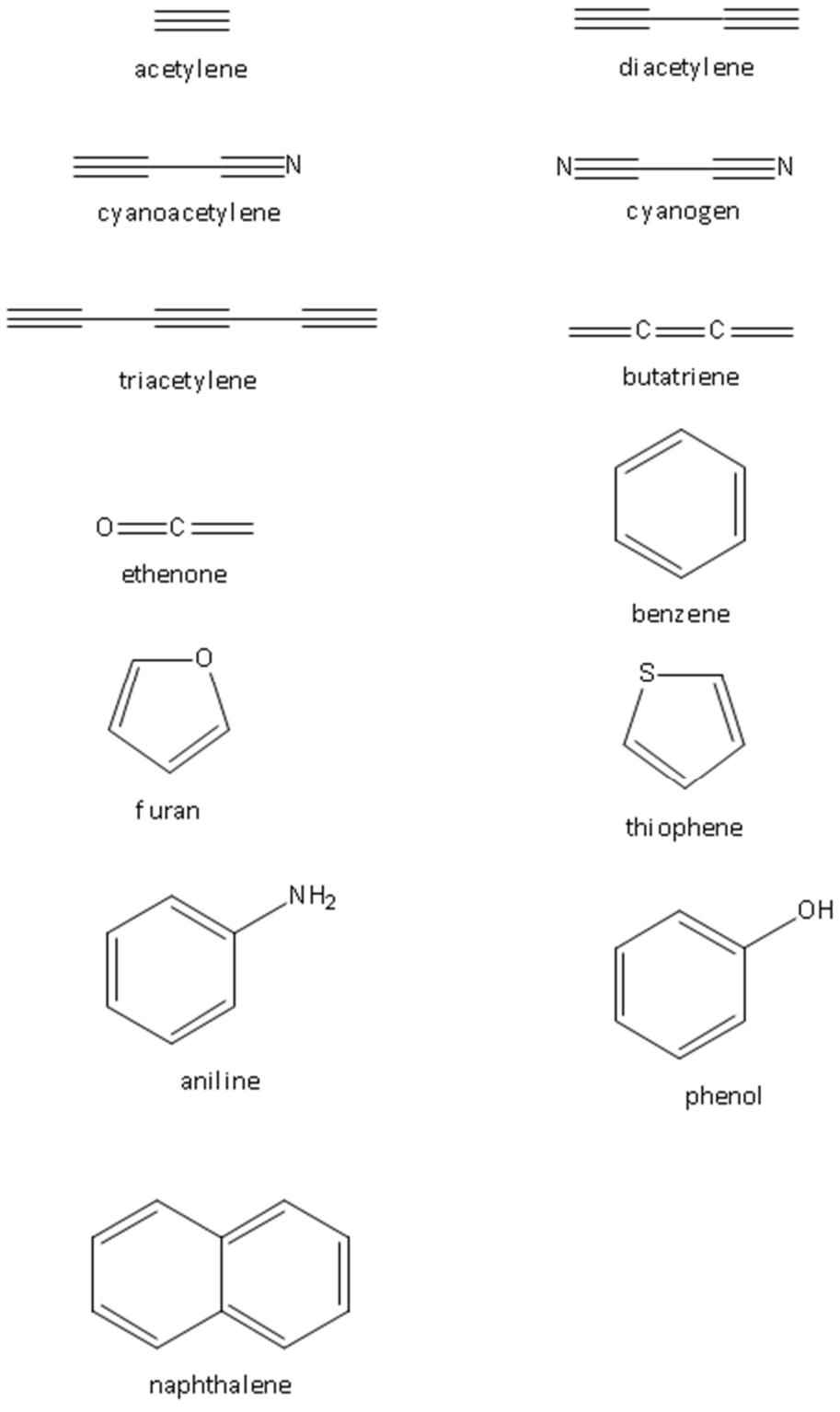


FIGURE 3.1: Systems explored in this work

TABLE 3.1: Summary of the basis sets examined in this work

All real frequencies for all systems studied				
STO-3g	ANO0 [9]			
cc-pVTZ [8]	ANO1 [9]			
aug-cc-pVTZ [13]	ANO2 [9]			
At least one imaginary frequency in a system studied				
3-21G	6-31+G [a]	6-31++G* [a]	6-311G	6-311+G** [a]
3-21+G	6-31++G [a]	6-31G**	6-311G* [a]	6-311++G** [a]
3-21++G	6-31G*	6-31+G**	6-311+G* [a]	aug-cc-pVDZ [a 13]
6-31G [a 12]	6-31+G*	6-31++G** [a]	6-311G** [a]	

[a] Imaginary frequencies for three or more systems

3.4 RESULTS AND DISCUSSION

The computations performed deliver a very clear answer. Planar, quasilinear, and linear molecules all experience this anomalous non-planarity/non-linearity problem when using Pople basis sets even at higher levels of theory. All molecules examined showed at least one imaginary/anomalously low frequency when set to their planar/linear stationary point for one or more Pople basis sets. This problem was highly exacerbated by diffuse functions, which tended to add imaginary frequencies. It was also clearly worsened in molecules with less rigidity such as triacetylene and diacetylene which exhibited many more imaginary frequencies. Polarization functions on the other hand seemed to help fix the problem and consistently lowered the number of imaginary frequencies computed, or even eliminated them entirely. It was also observed that as hypothesized, the use of Dunning and ANO basis sets represents the logical and easy fix to this problem. These basis sets showed no imaginary, nor anomalously low frequencies for nearly all of the molecules when at their planar/linear minima, evidencing that these were in fact their

true minima as previously and widely documented. The lone exception to proper behavior by correlation-consistent basis sets was aug-cc-pVDZ which produced imaginary frequencies in diacetylene, triacetylene, and butatriene. The cause of the inconsistency is hypothesized to be due to the basis set being too rich with *s* and *p* functions in comparison with higher level *d* and *f* functions.

One specific example is diacetylene computed at the CCSD/6-311G level of theory. In this case two imaginary frequencies, $\omega_7(\pi_g)$ and $\omega_9(\pi_u)$ at 1302*i* and 264*i* cm^{-1}, respectively, are found when the geometry is forced into $D_{\infty h}$ symmetry. A summary of the bending vibrational frequencies can be found in Table 3.2. The $\omega_7(\pi_g)$ imaginary normal mode was followed to find at minimum structure for the system at this level of theory and basis set. The minimum structure is depicted in Figure 3.2, while the $\omega_7(\pi_g)$ normal mode is represented in Figure 3.3. A harmonic frequency computation on this geometry confirms it is a minimum on the potential energy surface.

TABLE 3.2: Diacetylene CCSD/6-311G Bending Frequencies

Mode	Description	$\omega(cm^{-1})$
$\omega_6(\pi_g)$	C-C-H Bend	522
$\omega_7(\pi_g)$	C-C-C Bend	1302*i*
$\omega_8(\pi_u)$	C-C-H Bend	393
$\omega_9(\pi_u)$	C-C-C Bend	264*i*

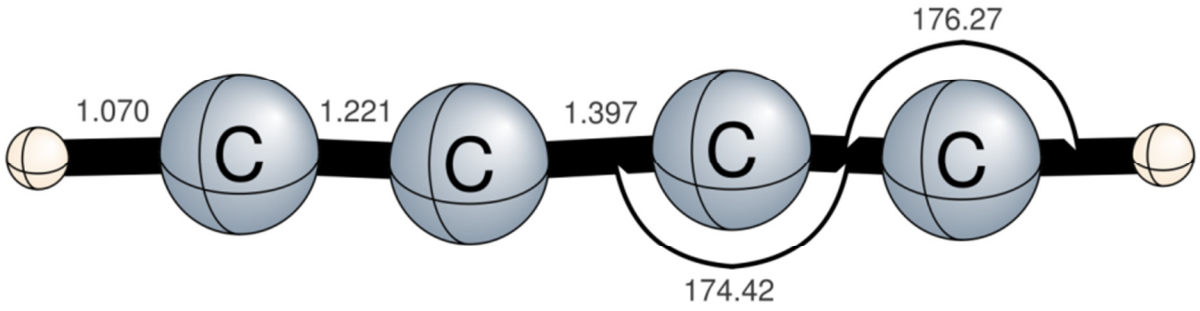


FIGURE 3.2: Diacetylene optimized geometry at the CCSD/6-311G level of theory.

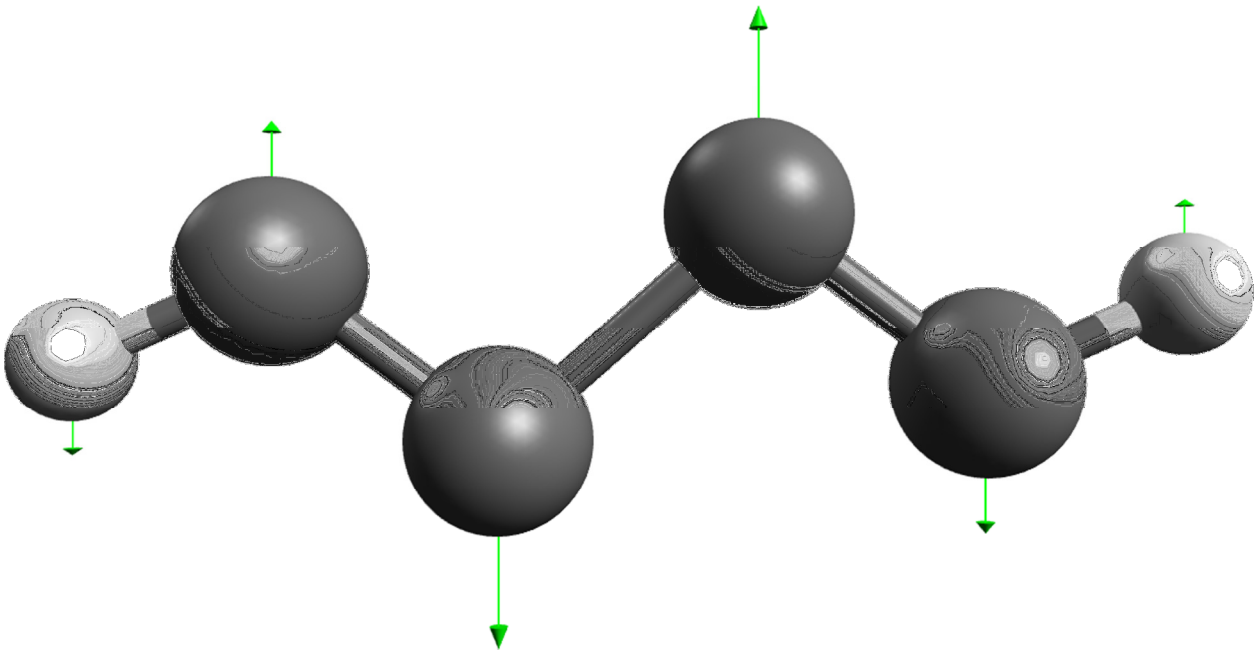


FIGURE 3.3: The $\omega_7(\pi_g)$ bending normal mode of diacetylene at 1302 cm^{-1}.

For further analysis, single-points were computed for diacetylene for progressive deviations from linearity along the imaginary frequency modes. These same computations were run for the cc-pVTZ basis set along the analogous $\omega_7(\pi_g)$ normal mode. This analysis comprises Figure 3.4. As can be seen, clear non-linear minima are predicted by the total energy computations for the 6-311G basis set but not by those for the cc-pVTZ basis set. Furthermore, the correct rediction at the HF level and the warped profile of the correlation energy plot for the 6-311G basis set suggest that the problem must lie within the correlation energy. In this case the 6-311G basis set is insufficient at spanning the space needed for an accurate description of the correlation energy.

The authors highly encourage the use of correlation consistent basis sets when computing properties of systems similar to the ones studied here. In our analysis, the ANO, cc-pVTZ, and aug-cc-pVTZ basis sets all lead to no imaginary frequencies for any of the thirteen systems investigated. However, though this moves in the correct direction, this recommendation is not full proof. Imaginary frequencies were found with the aug-cc-pVDZ basis set for multiple systems including triacetylene. The aug-cc-pVDZ basis set contains 4*s*3*p*2*d* functions for carbon and 3*s*2*p* functions for hydrogen. As seen in previous work,[1] saturating the basis set with lower angular momentum functions does not correct the problem. However, once higher angular momentum functions are included, such as *f* orbitals, only real frequencies are found for all systems as seen in the cc-pVTZ and aug-cc-pVTZ basis sets.

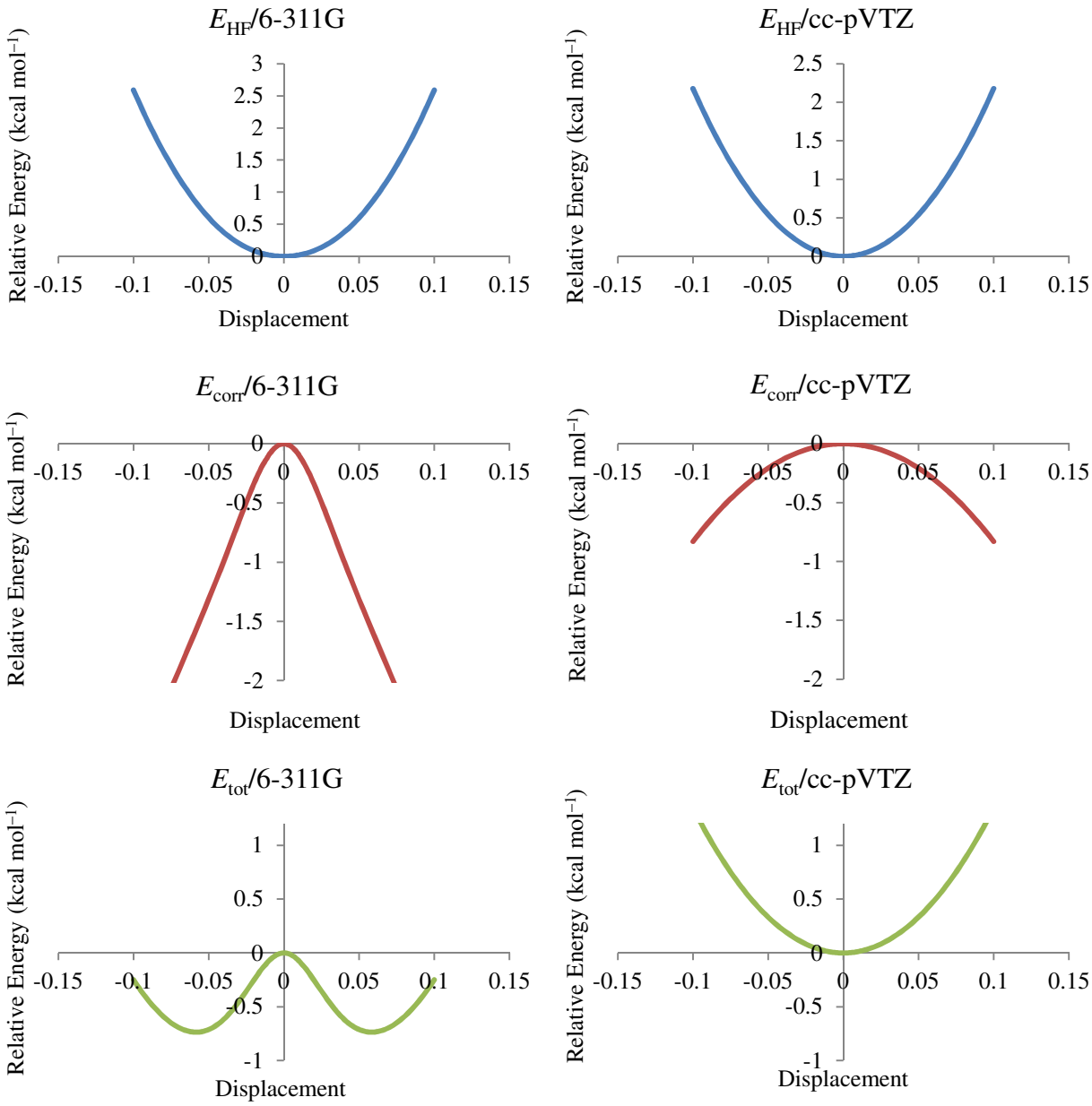


Figure 3.4: RHF energy, CCSD correlation energy, and CCSD total energies with 6-311G and cc-pVTZ basis sets along the mode of $\omega_7(\pi_g)$ in diacetylene.

3.5 CONCLUSIONS

Harmonic vibrational frequencies were computed for a selection of planar, linear, and quasilinear molecules with a large array of Pople, Dunning, and ANO basis sets at the MP2 and CC levels of theory. In general the use of unbalanced basis sets led to anomalous low or imaginary frequencies for all systems studied. Particularly, Pople basis sets without polarization functions do a very poor job at giving reasonable harmonic frequencies. Correlation-consistent basis sets, such as the Dunning or Atomic Natural Orbital series, correct for this error completely. An exception can be seen with the aug-cc-pVDZ basis. This shows that being too rich in *s* and *p* functions compared to *d* and *f* functions leads to anomalous results. In all we suggest never using Pople basis sets for second derivative properties with post-Hartree-Fock methods and instead opt for a more robust series.

3.6 REFERENCES

(1) Moran, D.; Simmonett, A. C.; Leach, F. E.; Allen, W. D.; Schleyer, P. V. R.; Schaefer, H. F. *J. Am. Chem. Soc.* **2006**, *128*, 9342–9343.

(2) Goodman, L.; Ozkabak, A. G.; Thakur, S. N. *J. Phys. Chem.* **1991**, *95*, 9044–9058.

(3) Martin, J. M. L.; Taylor, P. R.; Lee, T. J. *Chem. Phys. Lett.* **1997**, *275*, 414–422.

(4) Lampert, H.; Mikenda, W.; Karpfen, A. *J. Phys. Chem. A* **1997**, *5639*, 2254–2263.

(5) Dkhissi, A.; Adamowicz, L.; Maes, G. *J. Phys. Chem. A* **2000**, *104*, 2112–2119.

(6) Asturiol, D.; Duran, M.; Salvador, P. *J. Chem. Theory Comput.* **2009**, *5*, 2574–2581.

(7) Shahbazian, S. *Chem. Phys. Lett.* **2007**, *443*, 147–151.

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

(9) Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *86*, 4070.

(10) Martin, J. M. L.; Lee, T. J.; Taylor, P. R. *J. Chem. Phys.* **1998**, *108*, 676.

(11) Asturiol, D.; Duran, M.; Salvador, P. *J. Chem. Phys.* **2008**, *128*, 144108.

(12) Hehre, W. J. *J. Chem. Phys.* **1972**, *56*, 2257.

(13) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(14) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G.; Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J. *Curr. version, see http//www. cfour.* **2009**.

CHAPTER 4

INVESTIGATION OF 1-(1-DIAMANTYL) DIAMANTANE:

STABILIZING LONG ALKANE BONDS THROUGH INTRAMOLECULAR DISPERSION [†]

[†] D. B. Magers, A. E. Vaughn, J. M. Turney, and W. D. Allen. To be submitted to *Journal of the American Chemical Society*.

4.1 ABSTRACT

Dispersion is generally thought of as a qualitative description of an intermolecular force. In theoretical post-Hartree-Fock methods, dispersion is described by long-range electron correlation. In an electronic structure computation, the distinction between intramolecular versus intermolecular forces is unnecessary. Here the concept of intramolecular dispersion in 1-(1-diamantyl) diamantane is investigated by analyzing individual pair correlation energies with localized molecular orbitals. At the MP2/aug(H)-cc-pVDZ level of theory, 1-(1-diamanytl) diamantane is stabilized by intramolecular dispersion across the bridging C–C single bond by 35.7 kcal mol^{-1}. This result supports experimental evidence of the system's unexpected thermal stability despite obvious steric crowding. Unlike in previous work, we find that the H···H interactions across the central bridging bond are not the primary source of this stabilization. Instead, the intramolecular dispersion stabilization is attributed to general electron correlation found in the cages.

4.2 INTRODUCTION

Recent experimental work has synthesized one of the longest alkane bonds to date in diamondoid derivatives.[1,2] X-ray crystallography revealed an alkane C–C bond length of 1.647 Å in 1-(1-diamantyl) diamantane[1] (Figure 4.1) and 1.71 Å in 2-(1-diamantyl)[121]-tetramantane.[2] These are much longer than the C–C bond in ethane, the simplest alkane, of approximately 1.53 Å.[3,4] By cleverly constructing systems that possess steric crowding, long alkane bonds can be formed.

Steric crowding is a common method for elongating chemical bonds.[5] However, two problems arise. First, the strain caused by the steric crowding gets dispersed across multiple bonds and angles, making it difficult to localize the strain into a single bond.[5,6] Secondly, as the bond length increases, the bond dissociation energy (BDE) decreases.[7] The latter issue would cause one to believe that the bonded diamondoid cages are very unstable due to their long bond lengths. In contrast, 1-(1-diamantyl) diamantane has been shown to be quite thermally stable via differential scanning calorimetry with a melting point of 360 °C.[1] The unexpected stability of these diamondoid systems is attributed to "attractive dispersion interactions"[1] or "attractive steric interactions."[2] These intramolecular interactions take place across the strained alkane bond between the hydrogens. The hydrogens act as "dispersion energy donors"[8] compensating for the steric strain. Diamondoids are not the only systems that have this feature. The *t*-butyl groups in all *meta-tert*-butyl substituted hexaphenylethanes have been shown to be more stable than in its hexaphenylethane parent despite the additional steric crowding.[9]

The idea of H···H interactions being purely repulsive is incorrect as can be seen in the methane dimer.[10] This interaction curve looks like a standard Leonard-Jones potential, which means that the H···H interactions are attractive to the right of the minimum in the potential well. The question is if the well depth is significant. The design of 1-(1-diamantyl) diamantane and other similar diamondoids utilizes this attractive interaction to counterbalance the steric repulsion of the cages. This claim is defended by comparing density functional theory (DFT) results with dispersion corrected DFT. For 1-(1-diamantyl) diamantane, B3LYP with the 6-31G(d,p) basis set overestimates the experimental geometry by 0.027 Å and gives a BDE of 43.9 kcal mol^{-1}.[1] In contrast, B3LYP-D, B97-D, and M06-2X match the experimental geometry

better and give BDEs of 70.7, 64.5, and 65.8 kcal mol^{-1}, respectively. Also, dispersion corrected DFT better reproduces geometries of coupled diamondoids.[2]

The application of dispersion corrected DFT reveals that dispersion is crucial to understanding the nature of coupled diamondoids, as seen above. The use of *ab initio* methods would be preferred to analyze the nature of electron dispersion; however, we are limited due to the size of these systems. In the present work, we apply correlated *ab initio* methods to obtain the bond dissociation energy and provide a formal analysis of thermochemical dispersion for coupled diamondoids, allowing us to quantify the stabilization due to dispersion via analysis of pair correlation energies in 1-(1-diamantyl) diamantane.

4.3 FORMAL ANALYSIS OF DISPERSION

Dispersion can be thought of as instantaneous dynamical electron correlation, which is only a piece of the total electron correlation. In order to appropriately quantify dispersion and separate it from the other types of electron correlation, we provide the following formal analysis. The notation is as follows: Di (diamondoid monomer radical), A (generic hydrocarbon radical), E (total energy), E° (total energy without dispersion), E^{HF} (Hartree-Fock energy), ε (dispersion part of correlation energy), and $\varepsilon' = E - E^{HF} - \varepsilon$ (non-dispersion part of correlation energy).

For the bond dissociation of the diamondoid dimer,

$$Di_2 \rightarrow 2\ Di \qquad \text{(R1)}$$

we have

$$\Delta E_1 = 2\ E^\circ(Di) + 2\ \varepsilon(Di) - E^\circ(Di_2) - \varepsilon(Di_2) = 2\ E^\circ(Di) - E^\circ(Di_2) + \omega(Di_2) \qquad (4.1)$$

$$\omega(Di_2) = 2\ \varepsilon(Di) - \varepsilon(Di_2) = \text{dispersion stabilization of C–C bond in } Di_2 \qquad (4.2)$$

The target quantity of the investigation is thus given by

$$\omega(\mathrm{Di_2}) = \Delta E_1 + E^\circ(\mathrm{Di_2}) - 2\,E^\circ(\mathrm{Di}) \ . \tag{4.3}$$

In equation 4.1 we subdivide the total energy of reaction 1 (R1) into the total energy without dispersion (E°) and the dispersion part of the correlation energy (ε). By solving for $\omega(\mathrm{Di_2})$ we define the dispersion stabilization of the bridging C–C bond in $\mathrm{Di_2}$ as the difference of the dispersion part of the correlation in R1.

Now consider

$$2\ \mathrm{Di} + 2\ \mathrm{AH} \ \rightarrow\ 2\ \mathrm{DiH} + \mathrm{A_2}\ , \tag{R2}$$

whose reaction energy is

$$\Delta E_2 = 2\,E^\circ(\mathrm{DiH}) + E^\circ(\mathrm{A_2}) - 2\,E^\circ(\mathrm{AH}) - 2E^\circ(\mathrm{Di}) + \Delta\varepsilon_2\ , \tag{4.4}$$

where

$$\Delta\varepsilon_2 = 2\,\varepsilon(\mathrm{DiH}) + \varepsilon(\mathrm{A_2}) - 2\,\varepsilon(\mathrm{AH}) - 2\,\varepsilon(\mathrm{Di})\ . \tag{4.5}$$

In reaction R2, the reaction of the diamantane monomer radical (Di) with a generic hydrocarbon radical (AH) is given. Similar to before, the total energy of reaction can be subdivided into the total energy without dispersion (E°) and the dispersion part ($\Delta\varepsilon_2$).

From Eq. (4.4),

$$2\,E^\circ(\mathrm{Di}) = 2\,E^\circ(\mathrm{DiH}) + E^\circ(\mathrm{A_2}) - 2\,E^\circ(\mathrm{AH}) - \Delta E_2 + \Delta\varepsilon_2 \tag{4.6}$$

and from Eq. (4.3)

$$\omega(\mathrm{Di_2}) = \Delta E_1 + \Delta E_2 + E^\circ(\mathrm{Di_2}) + 2\,E^\circ(\mathrm{AH}) - E^\circ(\mathrm{A_2}) - 2\,E^\circ(\mathrm{DiH}) - \Delta\varepsilon_2\ . \tag{4.7}$$

Equation 4.6 is a simple rearrangement to solve for $2\,E^\circ(\mathrm{Di})$ in equation 4.4. Then this equation is substituted into Equation 4.3 to form equation 4.7. As a reminder, the $\omega(\mathrm{Di_2})$ value is the target quantity.

However, $\Delta E_1 + \Delta E_2$ equals the energy (ΔE_3) of the net reaction

$$Di_2 + 2\ AH \rightarrow 2\ DiH + A_2\ , \qquad \text{(R3)}$$

so that

$$\omega(Di_2) = \Delta E_3 - \Delta E_3^{HF} - \Delta\varepsilon_3' - \Delta\varepsilon_2 \qquad (4.8)$$

where ΔE_3^{HF} and $\Delta\varepsilon_3'$ are the Hartree-Fock and non-dispersion correlation contributions to the energy of reaction R3, respectively. By employing R3, $\omega(Di_2)$ has been simplified from equation 4.7 to equation 4.8.

Equation (4.8) is an exact expression with no approximations.

Reaction (R2) can be broken up into two steps:

$$2\ Di + 2\ AH \rightarrow 2\ DiH + 2\ A\ , \qquad \text{(R4)}$$

and

$$2\ A \rightarrow A_2\ . \qquad \text{(R5)}$$

Assumption (1):

Dispersion is negligible for the energy of the hydrogen-transfer reaction (R4).

Thus,

$$\Delta\varepsilon_2 \approx \varepsilon(A_2) - 2\varepsilon(A) = -\omega(A_2)\ , \qquad (4.9)$$

changing Eq. (8) to

$$\omega(Di_2) = \Delta E_3 - \Delta E_3^{HF} - \Delta\varepsilon_3' + \omega(A_2) \qquad (4.10)$$

This first assumption reduces the complexity of $\Delta\varepsilon_2$ in the current target equation 4.8.

Assumption (2):

Dispersion is negligible for the C–C bond dissociation energy of ethane.

Accordingly, in Eq. (4.10) choose A = CH_3 to obtain the working equation

$$\omega(\mathrm{Di}_2) = \Delta E_6 - \Delta E_6^{\mathrm{HF}} - \Delta\varepsilon_6' = \Delta E_6^{\mathrm{corr}} - \Delta\varepsilon_6' \qquad (4.11)$$

where the reaction of concern is

$$\mathrm{Di}_2 + 2\ \mathrm{CH}_4 \rightarrow 2\ \mathrm{DiH} + \mathrm{CH}_3\mathrm{CH}_3\ . \qquad (\mathrm{R6})$$

From this analysis we can see that to determine the stabilization energy due to dispersion in 1-(1-diamantyl) diamantane the total energy, the Hartree-Fock energy, and the non-dispersion part of the correlation energy of R6 must be found.

4.4 DISCUSSION

The optimized geometry of 1-(1-diamantyl) diamantane can be seen Figure 4.1. The central C–C bond of 1.646 Å is in excellent agreement with the experimental value of 1.647 Å,[1] in part due to error balancing. Noteworthy is the comparison of the experimental solid which exhibits a thermally averaged distance over the vibrational levels to the gas phase theoretical structure, r_e. By examining deviations in bond lengths in 2,2,3,3-tetramethylbutane in crystal versus gas phase geometries, the gas phase bond length of 1-(1-diamantyl) diamantane has been estimated by this empirical correction to be 1.655 Å.[2] We expect the bond length to decrease with increasing basis set size for the computational results, leading towards this empirical estimation.

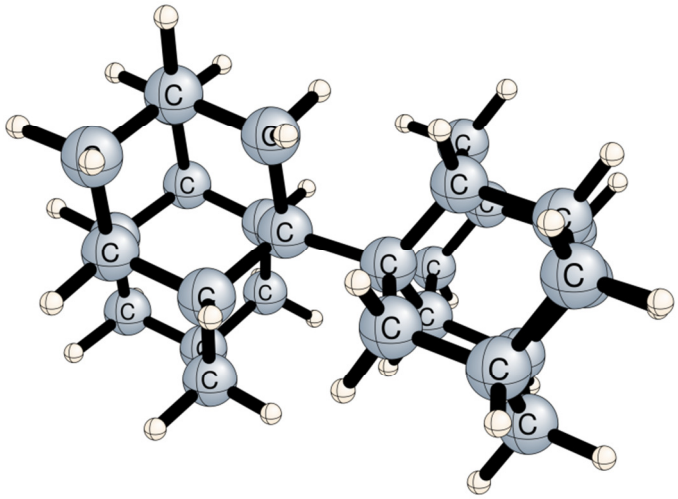


FIGURE 4.1: Optimized geometry of 1-(1-diamantyl) diamantane at the MP2/cc-pVDZ level of theory. The bridging C–C bond distance is 1.646 Å.

A few comments need to be made on the basis set choice. Since dispersion is a long range property, the use of a basis set with diffuse functions is important. Unfortunately, with the addition of diffuse functions comes increased computational cost. The hydrogen atoms form a shell around the globule system, and thus diffuse functions are only needed on the outside of the shell. We denote aug(H)-cc-pV*X*Z as the standard Dunning basis sets[11] with diffuse functions only added to hydrogens. In order to confirm this, energies of reaction were compared for propane plus methane yielding two ethanes (Table 4.1). Removing diffuse functions from the carbons has little effect on the energy of reaction with a difference of only 0.17 kcal mol^{-1} at CCSD(T). Since larger basis sets are not feasible for 1-(1-diamantyl) diamantane, we used the branched 2,2,3,3-tetramethylbutane, or hexamethylethane, as a benchmark to see the accuracy of the smaller basis aug(H)-cc-pVDZ (Table 4.2). We conclude that the aug(H)-cc-pVDZ basis set

TABLE 4.1: Energies of reaction of $C_3H_8 + CH_4 \rightarrow 2\ C_2H_6$ in kcal mol^{-1}

	HF	MP2	CCSD	CCSD(T)
aug(H)-cc-pVDZ[a]	0.86	2.41	1.92	2.14
aug-cc-pVDZ	0.87	2.58	2.06	2.31

[a] Augmented (diffuse) functions only added to hydrogens.

TABLE 4.2: Energies of reaction of $C_8H_{18} + 2\ CH_4 \rightarrow 2\ C_4H_{10} + C_2H_6$ in kcal mol^{-1}

	HF	MP2	CCSD	CCSD(T)
aug(H)-cc-pVDZ[a]	−10.68	5.90	1.52	3.69
aug(H)-cc-pVTZ[a]	−10.88	5.88	1.21	3.58

[a] Augmented (diffuse) functions only added to hydrogens.

will accurately predict the total correlation energy of $C_8H_{18} + 2\ CH_4 \rightarrow 2\ C_4H_{10} + C_2H_6$, and thus for R6, the reaction of concern. Following the notation in the dispersion analysis above, the total correlation energy is given as $\Delta E^{corr} = \Delta E - \Delta E^{HF}$. The total correlation energy of tetramethylbutane with the aug(H)-cc-pVDZ basis set shows a correlation stabilization of 14.37 kcal mol^{-1} while with the aug(H)-cc-VTZ basis set gives 14.46 kcal mol^{-1} at the CCSD(T) level of theory. This is excellent agreement, and we use this as a justification of the use of the aug(H)-cc-pVDZ basis set for analysis of 1-(1-diamantyl) diamantane.

Results for BDEs and C–C bond lengths of 1-(1-diamantyl) diamantane at varying levels of theory and basis sets can be seen in Table 4.3. As mentioned before, the B3LYP functional

TABLE 4.3: Bond lengths and bond dissociation energies with zero point vibration energy correction of 1-(1-diamantyl) diamantane with varying level of theories and basis sets

	BDE (kcal mol^{-1})	C–C (Å)
B3LYP/6-31G(d,p)[a]	43.9	1.674
B3LYP-D/6-31G(d,p)[a]	70.7	1.653
B3LYP-D3/6-31G(d,p)	50.5	1.670
B3LYP-CHG/6-31G(d,p)	59.2	1.665
B97D/6-31G(d,p)[a]	64.5	1.668
M06-2X/6-31G(d,p)[a]	65.8	1.648
B3PW91/6-31G(d,p)	46.1	1.660
MP2/cc-pVDZ[b]	75.63	1.646
CCSD(T)/aug(H)-cc-pVDZ[b]	71.25	—
Experiment[a]	—	1.647

[a] Results from Ref 1
[b] ZPVE correction with B3PW91/6-21(d,p)

without a dispersion correction performs poorly.[1,2] However, it is notable that the inclusion of different dispersion corrections (-D,[12] -D3,[13] and –CHG[14]) to B3LYP gives large differences in BDE and bond length. Benchmarking DFT for describing coupled diamondoids by analyzing rotation barriers has already been done.[2] The benchmark study concludes that the B3PW91 functional is the best at reproducing the physical properties of coupled diamondoids. This conclusion is in contrast with the BDEs shown in Table 4.3.

In order to divide the correlation energy into a dispersion and non-dispersion term as described in the analysis above, pair correlation energies (PCEs) are computed. PCEs are the correlation energies between any two electrons in a system. In MP2 and CCSD, the sum of the PCEs equals the total correlation at that level of theory and basis set. These values are automatically computed in electronic structure codes for both MP2 and CCSD. By dividing these pair energies into long and short range, dispersion and non-dispersion, categories we can quantify the dispersion in 1-(1-diamantyl) diamantane across the C–C bond. It is necessary to localize the canonical molecular orbitals in order to assign electron interactions as being either long or short range. In this analysis we have assigned any electron pair whose localized molecular orbitals reside on separate diamantane cages as long and all others as short. Referring back to equation 11, we can then evaluate the dispersion stabilization of the C–C bond in Di_2, $\omega(Di_2)$. At the MP2/aug(H)-cc-pVDZ level of theory, we find that 1-(1-diamantyl) diamantane is stabilized by dispersion by 35.7 kcal mol^{-1}. This dispersion stabilization is higher than the 26.8 kcal mol^{-1} predicted by previous work,[1] found by subtracting the B3LYP-D total energy from the B3LYP energy with a 6-31G(d,p) basis set in the bond dissociation reaction of 1-(1-diamantyl) diamantane.

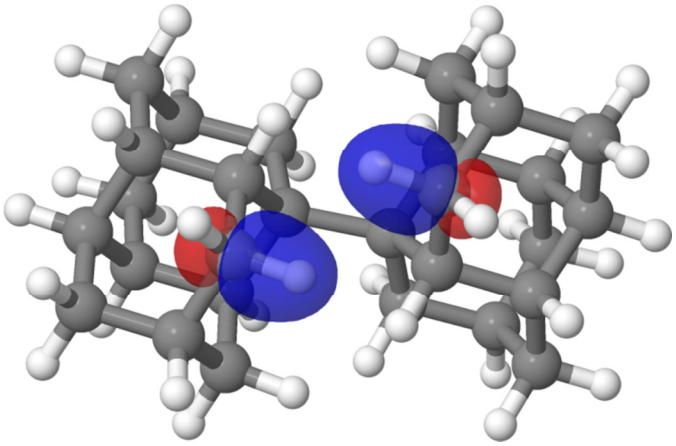
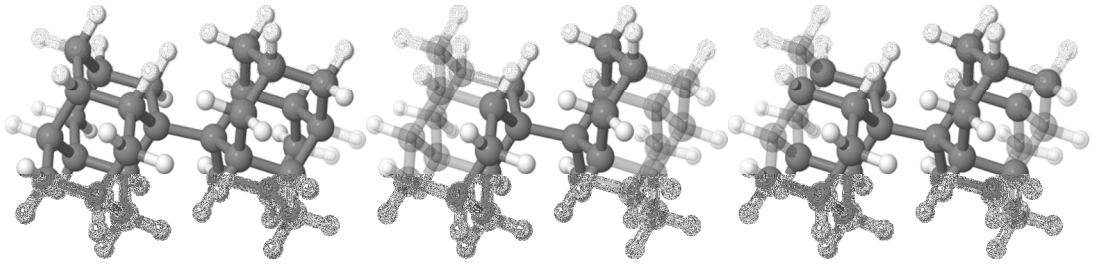


FIGURE 4.2: Two Boys localized molecular orbitals centered on C–H bonds present in the pair correlation energy analysis.

Quantifying the dispersion stabilization in 1-(1-diamantyl) diamantine is not the only information we can gather from the pair correlation energies. Since we have the correlation energy of every pair of electrons in localized orbitals in the system, we can determine whether the H···H interactions are indeed "dispersion energy donors"[8] and a significant contributor to the 35.7 kcal mol^{-1} dispersion stabilization. By summing only the PCEs from the inner H···H interactions across the central C–C bond, we find that these contribute only 21.4% of the 35.7 kcal mol^{-1} dispersion stabilization, or 7.6 kcal mol^{-1}. We partitioned the system into three physical divisions and computed the dispersion stabilization for only the interactions in each division. The divisions and the dispersion stabilization for each division can be seen in Figure 4.3 and Table 4.4, respectively. As expected, as the number of interactions is increased, the contribution to the total dispersion stabilization increases. From this analysis, it is apparent that the amount of correlation is highly dependent on distance, which does not come as a surprise. A plot of the pair correlation energy (a.u.) vs the pair distance is laid out in Figure 4.4. Since the

localized molecular orbitals appear as simple C–H and C–C bonds, the center of each bond was taken as the center of the orbital and thus the location for an electron. This plot not only shows the high correlation between distance and the correlation energy between two electrons, but additionally shows that the inner H···H interactions have no statistical significance compared to the other pairs and are indeed spread out among all the pairs.



Division One Division Two Division Three

FIGURE 4.3: Division of 1-(1-diamantyl) diamantine into three subgroups to analyze the dispersion stabilization (Refer to Table 4.4)

TABLE 4.4: Breakdown of the pair correlation energies based on physical divisions (Refer to Figure 4.3)

Subgroup of Dispersion Stabilization	Percent of Total Dispersion Stabilization (35.7 kcal mol^{-1})
Inner H---H Interactions	21.4
Division One Interactions	68.6
Division Two Interactions	85.4
Division Three Interactions	98.0

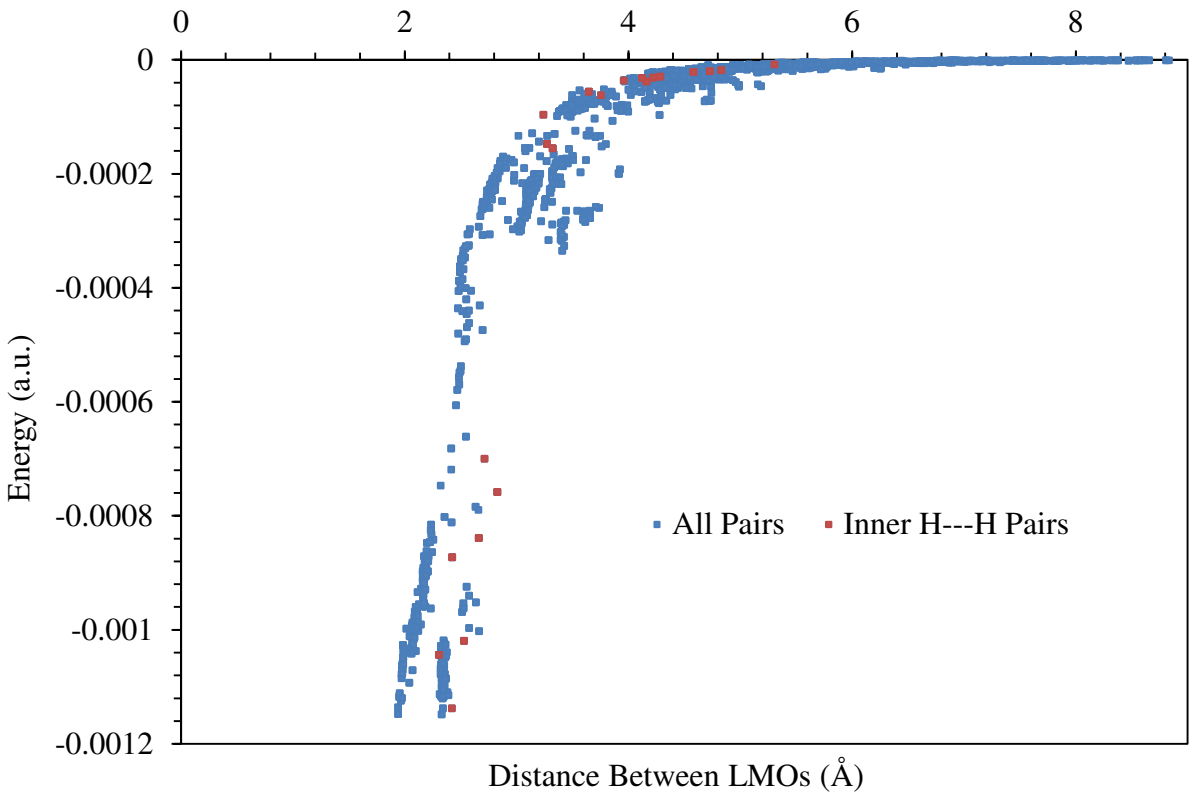


FIGURE 4.4: Plot of the pair correlation energy versus the distance of a pair of electrons. The location of an electron was taken as the center of either a C–H or C–C bond which corresponds to a localized molecular orbital.

4.5 COMPUTATIONAL METHODS

All *ab inito* computations were carried out with the MOLPRO program package.[15] Geometries were fully optimized at the MP2 level of theory with a cc-pVDZ basis set.[11] The open-shell molecules were optimized with an unrestrictued Hartree-Fock (UHF) reference wavefunction. Any spin contamination was always negligible. Single-point energies were computed at the

stated levels of theory and basis sets with the geometries mentioned. Open-shell molecule single points were computed with a restricted open-shell (ROHF) reference wavefunction. DFT computations were carried out with the Q-CHEM[16] and PSI4[17] program packages. The orbital localization and pair correlation energy analysis was conducted using the implementation in PSI4. The Pipek-Mezey[18] localization scheme was utilized.

4.6 CONCLUSIONS

1-(1-diamantyl) diamantane was analyzed using ab intio theoretical methods to quantify the stabilization due to dispersion across the central C–C bond. Following our formal definition of dispersion as described in the text, 1-(1-diamantyl) diamantine is stabilized by 35.7 kcal mol^{-1} at the MP2/aug(H)-cc-pVDZ level of theory. This energy stabilization is higher than predicted by Schreiner et al.[1] using DFT methods at 28.6 kcal mol^{-1}. We also find that the dispersion stabilization between any two electrons is highly dependent on the distance between them. As such the inner H···H interactions are not special and do not contribute any more significantly than other electron pairs.

4.7 REFERENCES

(1) Schreiner, P. R.; Chernish, L. V.; Gunchenko, P. A.; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Schlecht, S.; Dahl, J. E. P.; Carlson, R. M. K.; Fokin, A. A. *Nature* **2011**, *477*, 308–311.

(2) Fokin, A. A.; Chernish, L. V; Gunchenko, P. a; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Dahl, J. E. P.; Carlson, R. M. K.; Schreiner, P. R. *J. Am. Chem. Soc.* **2012**, *134*, 13641–13650.

(3) Duncan, J. L.; McKean, D. C.; Bruce, A. J. *J. Mol. Spectrosc.* **1979**, *74*, 361–374.

(4) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. *J. Chem. Soc. Perkin Trans. 2* **1987**, *2*, 1–19.

(5) De Silva, K. M. N.; Goodman, J. M. *J. Chem. Inf. Model.* **2005**, *45*, 81–87.

(6) Ruchardt, C.; Beckhaus, H.-D. *Angew. Chemie, Int. Ed.* **1985**, *24*, 529–616.

(7) Zavitsas, A. a. *J. Phys. Chem. A* **2003**, *107*, 897–898.

(8) Grimme, S.; Huenerbein, R.; Ehrlich, S. *Chemphyschem* **2011**, *12*, 1258–1261.

(9) Grimme, S.; Schreiner, P. R. *Angew. Chemie, Int. Ed.* **2011**, *50*, 12639–12642.

(10) Suzuki, T.; Takeda, T.; Kawai, H.; Fujiwara, K. *Pure Appl. Chem.* **2008**, *80*, 547–553.

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

(12) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(13) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(14) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(15) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 242–253.

(16) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V; Levchenko, S. V; O'Neill, D. P.; DiStasio, R. a; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. a; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. a; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. a; Rhee, Y.

M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

(17) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. a.; Fermann, J. T.; Mintz, B. J.; Burns, L. a.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. a.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556–565.

(18) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916.

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 CONCLUDING REMARKS

Through the use of high accuracy methods such as coupled cluster theory and advanced analysis techniques, subchemical accuracy can be obtained to within less than 1.0 kcal mol^{-1} for most thermochemical properties. With the rigorous design of coupled cluster, almost all of the electron correlation in a system can be recovered. The use of the correlation consistent basis sets series by Dunning and coworkers allows for a systematic progression to the complete basis set limit. Combining these two concepts together, the focal point analysis scheme provides a means to monitor the convergence of the correlation treatment and basis set to the exact energy. This analysis provides extremely high accuracy results without the brute force intractable computational task that would otherwise be required. Even with these advances in modern quantum chemistry and electronic structure theory, particular care must be made to ensure proper description of the system and property computed.

In Chapter 2, a detailed study was done on radical-radical abstraction reactions. These reactions are inherently multireference in nature and are poorly treated by standard single reference methods. In this case, the newer Mk-MRCC method was utilized to accurately form a wavefunction to describe the system. Using this method the focal point analysis technique was

used to extrapolate the energy of reaction potential energy surfaces to the complete basis set limit. These are the most accurate energy profiles of these systems to date.

Chapter 3 described a pitfall in utilizing insufficient basis sets to compute harmonic vibrational frequencies for certain classes of planar and linear systems involving π bonding. For a wide range of systems, anomalous results were computed with high level ab initio methods, such as MP2 and CCSD(T), due to poor basis sets. Particularly these unphysical results arise when high angular momentum functions are not included. The use of correlation consistent basis sets such as the Dunning or Atomic Natural Orbital series eliminate these problems.

In Chapter 4, the qualitative concept of dispersion is quantified in 1-(1-diamantyl) diamantane. Theoretically, dispersion can be considered long range dynamical electron correlation. By analyzing the pair correlation energies from an MP2 computation with localized molecular orbitals, the system was found to be stabilized by intramolecular dispersion despite the apparent steric repulsion. This unique analysis sheds light on a class of molecules that are highly strained, yet thermally stable.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 2 [†]

[†] Chia-Hua Wu,[‡] D. Brandon Magers,[‡] Lawrence B. Harding, Stephen J. Klippenstein, and Wesley D. Allen. To be submitted to *Journal of Chemical Theory and Computation*.

[‡] These authors contributed equally.

A.1: PRELIMINARY ANALYSIS

A.1.1: VALIDITY OF TIKHONOV QUADRATIC ENERGY EXTRAPOLATION

TABLE A.1: Mk-MRCCSD/cc–pVTZ and Mk-MRCCSD(T)/cc–pVTZ energies from Tikhonov quadratic extrapolation vs. explicit computation for $H + C_2H_5 \rightarrow H_2 + C_2H_4$

Mk–MRCCSD/cc–pVTZ

		$E(\omega)$		$E(\omega=0)$	$E(\omega=0)$
S (Å)	ω=0.001	ω=0.002	ω=0.003	extrapolated	explicit
0.0	–79.508353301	–79.508353101	–79.508352763	–79.508353369	–79.508353367
0.1	–79.503446346	–79.503446147	–79.503445815	–79.503446412	–79.503446412
0.2	–79.499432548	–79.499432339	–79.499431987	–79.499432619	–79.499432618
0.3	–79.496759668	–79.496759452	–79.496759081	–79.496759743	–79.496759740
0.4	–79.494877096	–79.494876872	–79.494876492	–79.494877173	–79.494877171
0.5	–79.493696343	–79.493696115	–79.493695732	–79.493696420	–79.493696420
0.9	–79.491354977	–79.491354716	–79.491354281	–79.491355064	–79.491355064
1.0	–79.491091690	–79.491091423	–79.491090972	–79.491091781	–79.491091778
1.2	–79.490778500	–79.490778225	–79.490776967	–79.490778820	–79.490778591
1.4	–79.490577950	–79.490577669	–79.490577192	–79.490578046	–79.490578047

Mk–MRCCSD(T)/cc–pVTZ

		$E(\omega)$		$E(\omega=0)$	$E(\omega=0)$
S (Å)	ω=0.001	ω=0.002	ω=0.003	extrapolated	explicit
0.0	–79.521032489	–79.521032418	–79.521032204	–79.521032490	–79.521032489
0.1	–79.516042142	–79.516042071	–79.516041857	–79.516042142	–79.516042142
0.2	–79.512407554	–79.512407480	–79.512407256	–79.512407555	–79.512407554
0.3	–79.509517882	–79.509517805	–79.509517576	–79.509517886	–79.509517882
0.4	–79.507458761	–79.507458682	–79.507458444	–79.507458763	–79.507458761
0.5	–79.505973263	–79.505973182	–79.505972940	–79.505973262	–79.505973263
0.9	–79.503503552	–79.503503461	–79.503503188	–79.503503554	–79.503503552
1.0	–79.503250117	–79.503250024	–79.503249743	–79.503250118	–79.503250117
1.2	–79.502899447	–79.502899407	–79.502899117	–79.502899453	–79.502899447
1.4	–79.502677420	–79.502677302	–79.502677026	–79.502677402	–79.502677420

A.1.2: LEADING CI COEFFICIENTS IN CASSCF WAVE FUNCTIONS

TABLE A.2: Leading CI coefficients (C_i) in CASSCF wave functions along the entrance channel of hydrogen abstraction reactions

$CH_3 + C_2H_5 \rightarrow CH_4 + C_2H_4$

	CAS(2,2)/cc-pVTZ		CAS(4,4)/cc-pVTZ			
S (Å)	C_1	C_2	C_1	C_2	C_3	C_4
–3.0	0.978	–0.207	0.974	–0.206	–0.090	< 0.050
–1.8	0.978	–0.206	0.974	–0.206	–0.089	< 0.050
–0.6	0.978	–0.208	0.973	–0.206	–0.088	< 0.050
–0.2	0.976	–0.217	0.968	–0.213	–0.076	< 0.050
0.0	0.972	–0.236	0.960	–0.240	–0.065	–0.063
0.2	0.953	–0.303	0.941	–0.301	–0.074	0.074
0.6	0.866	–0.499	0.862	–0.490	–0.071	0.051
1.0	0.793	–0.609	0.792	–0.603	–0.063	< 0.050
1.4	0.758	–0.653	0.756	–0.648	–0.062	0.054
1.8	0.738	–0.675	0.735	–0.672	–0.061	0.056
2.2	0.726	–0.688	0.723	–0.684	–0.060	0.057
2.6	0.714	–0.700	0.711	–0.696	–0.060	0.058
3.0	0.713	–0.701	0.710	–0.698	–0.059	0.059

$H + C_2H_5 \rightarrow H_2 + C_2H_4$

	CAS(2,2)/cc-pVTZ		CAS(4,4)/cc-pVTZ			
S (Å)	C_1	C_2	C_1	C_2	C_3	C_4
–2.0	0.978	–0.207	0.978	–0.175	–0.057	–0.056
–1.8	0.978	–0.207	0.978	–0.175	–0.057	–0.056
–1.6	0.978	–0.207	0.973	–0.205	–0.105	< 0.05
–1.4	0.978	–0.207	0.972	–0.205	–0.104	< 0.05
–1.2	0.978	–0.207	0.978	–0.175	–0.057	–0.056
–1.0	0.978	–0.208	0.972	–0.205	–0.100	< 0.05
–0.8	0.978	–0.211	0.970	–0.206	–0.093	< 0.05
–0.6	0.976	–0.217	0.967	–0.212	–0.080	< 0.05
–0.4	0.972	–0.236	0.958	–0.242	–0.069	–0.069
–0.2	0.952	–0.305	0.939	–0.304	0.077	0.077
0.0	0.919	–0.394	0.910	–0.386	–0.071	–0.071
0.2	0.876	–0.483	0.871	–0.474	–0.071	–0.055
0.4	0.834	–0.551	0.831	–0.545	–0.072	< 0.05
0.6	0.802	–0.597	0.798	–0.593	–0.071	0.053
0.8	0.779	–0.627	0.775	–0.624	–0.069	0.056
1.0	0.761	–0.648	0.757	–0.646	–0.068	0.059
1.2	0.748	–0.663	0.744	–0.661	–0.067	0.060
1.4	0.739	–0.674	0.735	–0.671	–0.067	0.061
1.6	0.731	–0.682	0.727	–0.679	–0.066	0.062
1.8	0.726	–0.688	0.722	–0.685	–0.066	0.062
2.0	0.721	–0.693	0.718	–0.690	–0.065	0.063
5.0	0.707	–0.707	0.673	–0.673	–0.207	0.207
10.0	0.707	–0.707	–0.707	0.707	< 0.05	< 0.05

$H + HCO \rightarrow H_2 + CO$

	CAS(2,2)/cc-pVQZ		CAS(4,4)/cc-pVQZ			
S (Å)	C_1	C_2	C_1	C_2	C_3	C_4
–1.6	0.989	–0.145	0.977	–0.112	–0.111	0.077
–1.0	0.990	–0.143	0.977	–0.113	–0.110	0.077
–0.7	0.990	–0.141	0.977	–0.114	–0.108	0.077
–0.2	0.976	–0.216	0.963	–0.217	–0.156	< 0.05
0.2	0.933	–0.359	0.919	–0.355	–0.157	0.061
0.6	0.866	–0.501	0.851	–0.493	–0.151	0.088
1.0	0.803	–0.596	0.789	–0.586	–0.143	0.107
1.4	0.763	–0.647	0.749	–0.636	–0.138	0.117
1.8	0.739	–0.673	0.726	–0.662	–0.134	0.122
2.4	0.721	–0.693	0.708	–0.681	–0.131	0.126
3.0	0.713	–0.701	0.700	–0.689	–0.129	0.127

A.2: RADICAL-RADICAL HYDROGEN ABSTRACTION REACTION PATHS

A.2.1: $H + BeH \rightarrow H_2 + Be$

TABLE A.3: Energy (kcal mol^{-1}) relative to separated reactants along the H + BeH reaction path for different levels of theory with the cc-pVTZ basis set[a]

S(Å)	CAS(2,2)	MRCI	MRCI+Q	CASPT2	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	FCI
–5.0	–45.18	–58.46	–60.46	–54.21	–53.48	–60.19	–59.85	–59.65
–3.0	–45.05	–58.49	–60.49	–54.23	–53.48	–60.20	–59.88	–59.69
–2.4	–44.44	–58.22	–60.27	–53.93	–53.14	–59.91	–59.63	–59.46
–2.0	–43.08	–57.46	–59.59	–53.09	–52.24	–59.14	–58.90	–58.77
–1.8	–41.78	–56.66	–58.86	–52.21	–51.33	–58.33	–58.09	–58.02
–1.6	–39.80	–55.37	–57.68	–50.81	–49.88	–57.05	–56.88	–56.80
–1.4	–36.87	–53.38	–55.83	–48.66	–47.69	–55.02	–54.93	–54.90
–1.2	–32.71	–50.40	–53.06	–45.52	–44.54	–52.06	–52.03	–52.06
–1.0	–27.04	–46.11	–48.99	–41.11	–40.19	–47.77	–47.80	–47.92
–0.8	–19.75	–40.11	–43.17	–35.27	–34.61	–41.60	–41.78	–42.05
–0.5	–8.03	–27.61	–30.61	–23.88	–22.66	–28.08	–28.65	–29.34
–0.3	–1.92	–18.29	–20.70	–15.66	–13.69	–18.56	–19.03	–19.73
–0.1	1.35	–10.79	–12.49	–9.12	–6.94	–10.95	–11.29	–11.88
0.0	2.03	–8.05	–9.42	–6.72	–4.67	–8.18	–8.46	–8.97
0.1	2.26	–5.93	–7.02	–4.88	–3.04	–6.03	–6.27	–6.69
0.2	2.23	–4.33	–5.18	–3.51	–1.93	–4.43	–4.60	–4.96
0.3	2.04	–3.15	–3.81	–2.52	–1.21	–3.24	–3.37	–3.66
0.4	1.77	–2.30	–2.80	–1.82	–0.73	–2.38	–2.48	–2.70
0.6	1.22	–1.24	–1.52	–0.97	–0.28	–1.28	–1.34	–1.49
0.8	0.77	–0.69	–0.85	–0.55	–0.11	–0.73	–0.77	–0.85
1.0	0.46	–0.41	–0.50	–0.34	–0.07	–0.44	–0.46	–0.50
1.2	0.26	–0.26	–0.31	–0.22	–0.06	–0.28	–0.29	–0.32
1.4	0.14	–0.18	–0.21	–0.16	–0.06	–0.19	–0.19	–0.21
4.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[a]The geometries along the reaction path are optimized at the CASPT2/cc–pVTZ level of theory; the geometry at S = 8 Å is taken as the asymptotic limit.

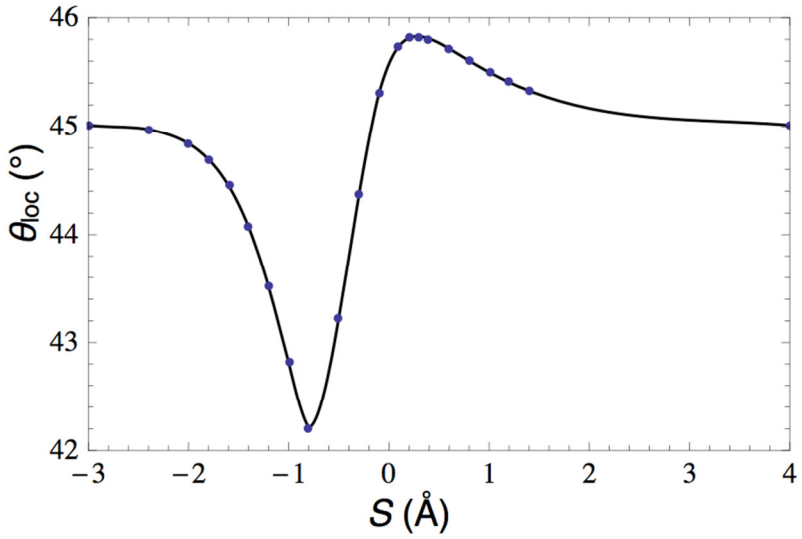


FIGURE A.1: Boys orbital localization angle θ_{loc} along the H + BeH reaction path.

A.2.2: $H + NH_2 \rightarrow H_2 + NH$

TABLE A.4: Energy (kcal mol^{-1}) relative to separated reactants along the $H + NH_2$ reaction path for different levels of theory with the 6-31G* basis set[a]

S(Å)	ROHF	MRCI	MRCI+Q	CASPT2	Mk–MRPT2	Mk–MRCCSD	Mk–MRCCSD(T)	FCI
–2.4	26.19	29.41	29.41	31.56	33.21	28.75	29.77	29.95
–2.2	26.23	29.42	29.42	31.56	33.21	28.76	29.77	29.95
–2.0	26.38	29.53	29.52	31.65	33.30	28.87	29.87	30.05
–1.8	26.71	29.81	29.81	31.91	33.56	29.16	30.14	30.33
–1.6	27.44	30.41	30.38	32.47	34.14	29.75	30.72	30.90
–1.4	28.75	31.37	31.25	33.31	35.10	30.67	31.60	31.75
–1.2	31.01	33.00	32.72	34.80	36.73	32.21	33.06	33.18
–0.9	37.36	37.12	36.13	38.60	40.89	35.92	36.48	36.48
–0.8	40.40	38.73	37.27	40.03	42.51	37.26	37.65	37.58
–0.7	43.77	40.13	38.06	41.09	43.76	38.32	38.50	38.34
–0.6	47.14	40.90	38.23	41.03	43.83	38.80	38.71	38.50
–0.5	49.48	40.29	37.50	38.51	41.23	38.22	37.83	37.70
–0.4	48.62	37.73	35.23	34.08	37.08	35.64	35.01	35.27
–0.2	37.51	30.37	28.58	28.15	31.22	28.45	28.02	28.42
0.0	25.52	21.50	20.53	20.11	22.27	20.34	20.11	20.34
0.2	16.11	13.97	13.51	13.08	14.46	13.37	13.26	13.37
0.4	9.66	8.56	8.37	8.00	8.85	8.27	8.23	8.28
0.6	5.58	5.03	4.96	4.69	5.20	4.89	4.89	4.91
0.8	3.14	2.86	2.84	2.66	2.96	2.80	2.80	2.81
1.0	1.73	1.58	1.57	1.46	1.64	1.55	1.56	1.56
1.2	0.93	0.84	0.84	0.77	0.88	0.82	0.84	0.83
1.4	0.48	0.42	0.43	0.39	0.45	0.42	0.43	0.42
1.8	0.11	0.09	0.09	0.07	0.10	0.08	0.10	0.08
2.2	0.02	0.01	0.00	0.00	0.01	0.00	0.02	0.00

[a]The geometries along the reaction path are optimized at the CASPT2/6-31G* level of theory; the geometry at S = 10 Å is taken as the asymptotic limit.

A.2.3: $CH_3 + C_2H_5 \rightarrow CH_4 + C_2H_4$

TABLE A.5: Energy (kcal mol^{-1}) relative to separated reactants along the $CH_3 + C_2H_5$ reaction path for different levels of theory at the CBS limit[a]

S(Å)	CAS(2,2)	MRCI	MRCI+Q	CASPT2	Mk-MRCCSD	Mk-MRCCSD(T)
0.2	3.51	–10.88	–15.31	–18.66	–16.69	–18.59
0.4	8.31	–2.96	–6.59	–10.31	–8.19	–9.75
0.6	7.90	0.31	–2.14	–5.06	–3.51	–4.65
0.8	5.63	0.74	–0.81	–2.72	–1.68	–2.47
1.0	3.54	0.32	–0.68	–1.89	–1.18	–1.71
1.2	2.06	–0.11	–0.77	–1.55	–1.05	–1.42
1.8	0.30	–0.42	–0.64	–0.88	–0.61	–0.73
2.4	0.34	–0.24	–0.41	–0.50	–0.39	–0.50
2.6	0.35	–0.23	–0.40	–0.47	–0.40	–0.51

[a]The geometries along the reaction path are optimized at the CASPT2/cc–pVQZ level of theory; the geometry at S = 10 Å is taken as the asymptotic limit.

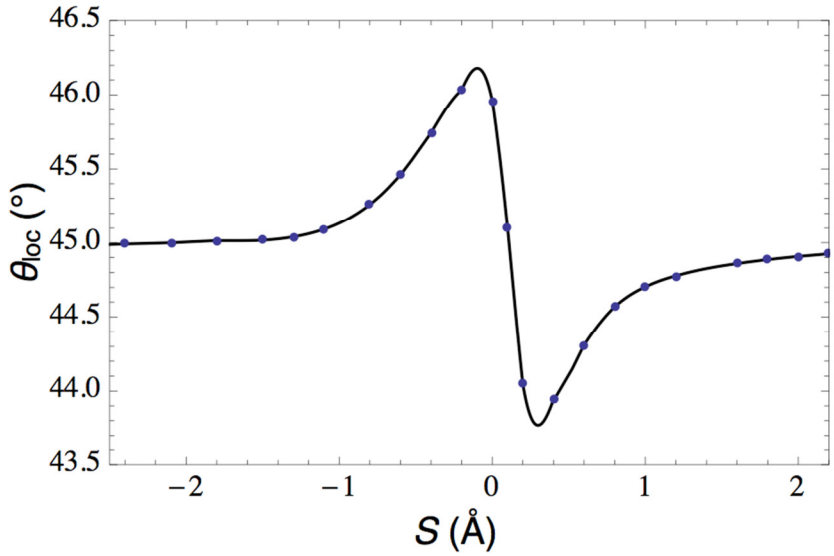


FIGURE A.2: Boys orbital localization angle θ_{loc} along the $CH_3 + C_2H_5$ reaction path

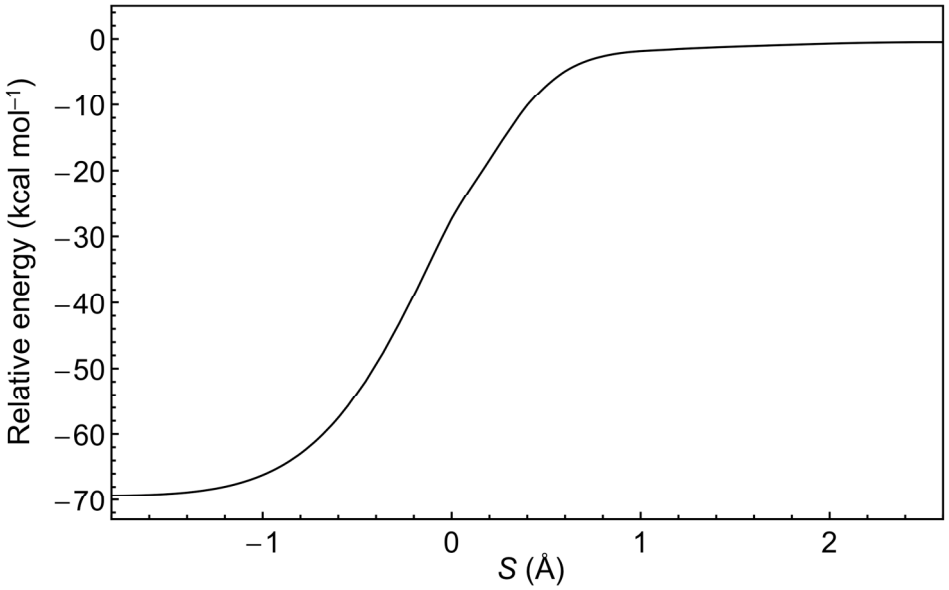


FIGURE A.3: The one-dimensional potential energy curve for CH_3 + C_2H_5 in CASPT2 level of theory.

TABLE A.6: Projected vibrational frequencies (ω, cm^{-1}) along the $CH_3 + C_2H_5$ reaction path

Reaction coordinate S(Å)									
0.2	0.4	0.6	0.8	1.0	1.2	1.8	2.4	2.6	10
44	42	34	23	14	6	9*i*	25*i*	18*i*	11*i*
205	200	159	124	96	73	28	31	22	5*i*
262	237	204	170	135	104	38	39	42	5*i*
581	457	464	366	282	216	94	44	50	6
592	547	515	406	308	230	94	75	59	10
765	608	597	488	370	281	151	112	109	108
797	751	631	643	590	559	523	515	514	513
905	796	722	648	610	587	574	568	567	565
952	874	835	823	819	818	817	820	820	818
1058	1068	1045	1010	992	988	988	993	993	990
1171	1120	1077	1119	1114	1105	1093	1092	1092	1091
1242	1233	1226	1220	1215	1212	1208	1208	1208	1207
1290	1281	1263	1300	1362	1389	1407	1409	1409	1410
1339	1369	1415	1437	1437	1437	1438	1438	1439	1439
1458	1450	1442	1438	1437	1437	1439	1439	1439	1439
1477	1458	1447	1458	1475	1479	1487	1489	1489	1490
1483	1479	1474	1472	1479	1490	1501	1502	1503	1503
1543	1542	1529	1518	1511	1508	1505	1504	1504	1505
1670	1830	1995	2301	2606	2801	3005	3035	3034	3037
3100	3116	3126	3115	3111	3110	3116	3119	3119	3123
3151	3140	3139	3156	3164	3168	3170	3167	3167	3171
3195	3193	3195	3187	3178	3174	3171	3171	3171	3172
3240	3225	3204	3198	3201	3204	3207	3209	3209	3209
3243	3274	3299	3304	3309	3313	3318	3321	3321	3321
3244	3275	3311	3339	3354	3359	3363	3363	3363	3365
3297	3296	3313	3341	3355	3360	3363	3364	3365	3365

TABLE A.7: Electronic energies (E_h) for CBS extrapolation at selected points along the CH_3 + C_2H_5 reaction path with CASPT2/cc–pVQZ optimized geometries

S = 0.2 Å

	CAS(2,2)	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–118.1909431	–118.7717974	–118.7912753	–118.7317432	–118.6945237	–118.7707215
cc-pVQZ	–118.1987700	–118.8054959	–118.8270034	–118.7720646	–118.7247390	–118.8041376
cc-pV5Z	–118.2007577					

S = 1.0 Å

	CAS(2,2)	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–118.1908535	–118.7480461	–118.7656197	–118.7057141	–118.6773170	–118.7482905
cc-pVQZ	–118.1986703	–118.7811540	–118.8006055	–118.7455615	–118.7071352	–118.7811723
cc-pV5Z	–118.2006799					

S = 10.0 Å

	CAS(2,2)	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–118.1962756	–118.7464266	–118.7632647	–118.7029491	–118.6779038	–118.7473796
cc-pVQZ	–118.2042263	–118.7793909	–118.7980397	–118.7426598	–118.7076817	–118.7801630
cc-pV5Z	–118.2062941					

TABLE A.8: Relative energies (ΔE, kcal mol^{-1}) at selected points along the $CH_3 + C_2H_5$ reaction path for each level of theory and basis sets leading to the CBS limit[a]

S = 0.2 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	3.35	–15.92	–17.58	–18.07	–10.43	–14.65
cc-pVQZ	3.42	–16.38	–18.17	–18.45	–10.70	–15.04
cc-pV5Z	3.47	[–16.52]	[–18.37]	[–18.57]	[–10.78]	[–15.16]
CBS	[+3.51]	[–16.69]	[–18.59]	[–18.71]	[–10.88]	[–15.31]

S = 1.0 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	3.40	–1.02	–1.48	–1.74	0.37	–0.57
cc-pVQZ	3.49	–1.11	–1.61	–1.82	0.34	–0.63
cc-pV5Z	3.52	[–1.13]	[–1.65]	[–1.84]	[+0.34]	[–0.65]
CBS	[+3.54]	[–1.18]	[–1.71]	[–1.89]	[+0.32]	[–0.68]

[a] The reference point is the energy of the separated reactants, taken from computations at S = 10.0 Å. CBS limits were evaluated using eqs 12 and 13 of the text.

A.2.4: $H + C_2H_5 \rightarrow H_2 + C_2H_4$

TABLE A.9: Energy (kcal mol^{-1}) relative to separated reactants along the $H + C_2H_5$ reaction path for different levels of theory at the CBS limit[a]

S(Å)	CAS(2,2)	MRCI	MRCI+Q	CASPT2	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)
0.0	2.77	–8.06	–11.19	–12.28	–11.24	–12.27	–12.75
0.1	3.68	–5.39	–8.05	–9.26	–7.70	–9.12	–9.53
0.2	3.86	–3.51	–5.69	–6.85	–5.15	–6.49	–7.19
0.3	3.59	–2.26	–4.00	–5.02	–3.38	–4.74	–5.28
0.4	3.10	–1.48	–2.85	–3.69	–2.18	–3.49	–3.90
0.5	2.56	–1.01	–2.08	–2.75	–1.48	–2.55	–2.94
0.6	2.04	–0.74	–1.57	–2.10	–1.00	–2.06	–2.20
0.8	1.22	–0.48	–0.99	–1.31	–0.59	–1.24	–1.38
0.9	0.92	–0.42	–0.82	–1.07	–0.50	–1.00	–1.12
1.0	0.69	–0.38	–0.70	–0.88	–0.43	–0.84	–0.93
1.2	0.37	–0.31	–0.52	–0.62	–0.33	–0.60	–0.65

[a]The geometries along the reaction path are optimized at the CASPT2/cc-pVQZ level of theory; the geometry at S = 10 Å is taken as the asymptotic limit.

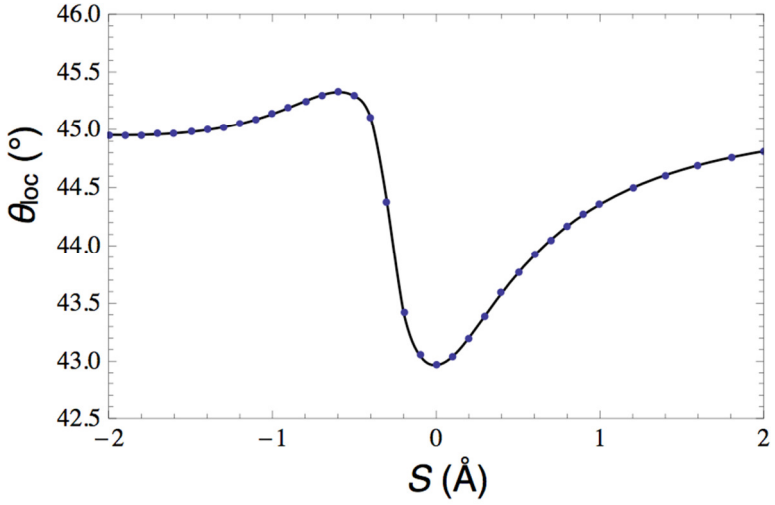


FIGURE A.4: Boys orbital localization angle θ_{loc} along the $H + C_2H_5$ reaction path

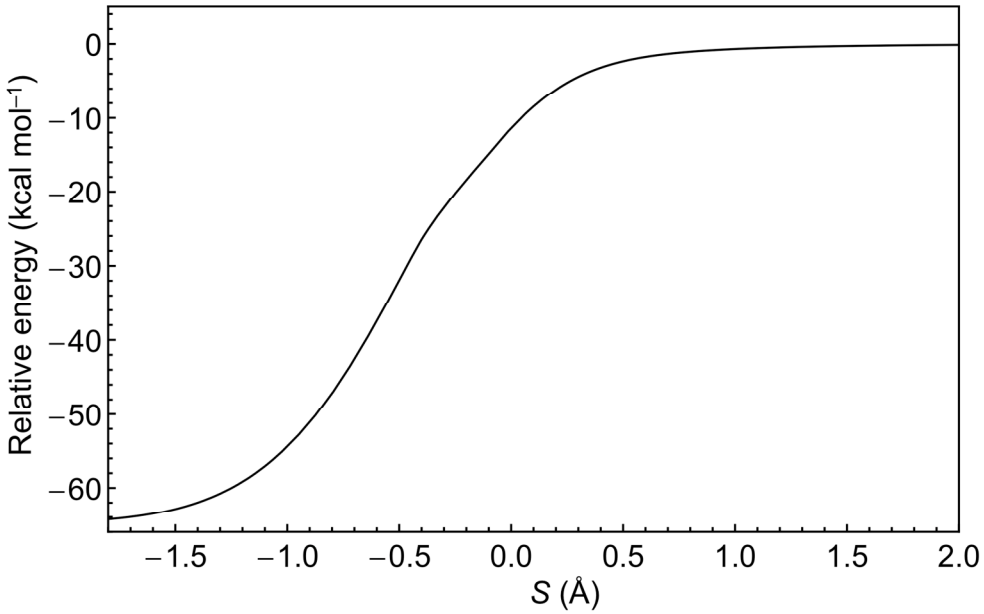


FIGURE A.5: The one-dimensional potential energy curve for H + C_2H_5 in CASPT2 level of theory.

TABLE A.10: Projected vibrational frequencies (ω, cm^{-1}) along the H + C_2H_5 reaction path

Reaction coordinate S(Å)								
0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
459	427	388	344	300	259	221	189	160
492	475	434	378	322	272	230	194	164
681	620	555	490	430	377	331	292	258
794	760	733	706	678	650	624	601	582
841	832	826	822	820	818	818	817	817
901	927	965	992	998	995	990	988	986
1217	1216	1216	1217	1182	1150	1131	1119	1111
1229	1218	1218	1232	1216	1215	1214	1212	1211
1340	1366	1319	1242	1274	1311	1340	1361	1376
1475	1412	1391	1415	1436	1452	1465	1474	1478
1488	1479	1477	1475	1474	1474	1475	1476	1482
1552	1552	1540	1531	1524	1518	1515	1512	1510
1750	1839	1982	2150	2324	2483	2616	2723	2804
3149	3144	3139	3133	3128	3123	3120	3118	3116
3195	3196	3197	3198	3200	3195	3190	3185	3182
3236	3228	3219	3210	3202	3201	3202	3204	3205
3299	3301	3302	3305	3307	3309	3311	3313	3315

Reaction coordinate S(Å)								
0.9	1	1.2	1.4	1.6	1.8	2	5	10
136	115	82	58	41	28	19	7*i*	5i
139	118	86	63	47	35	27	5*i*	5*i*
230	206	170	146	131	121	115	107	107
567	554	538	528	522	518	516	513	513
817	817	817	817	817	817	817	818	818
986	986	987	988	989	989	990	990	990
1105	1101	1097	1094	1093	1092	1091	1091	1091
1210	1210	1209	1208	1208	1207	1207	1207	1207
1386	1392	1401	1405	1407	1408	1409	1410	1410
1480	1482	1485	1487	1488	1489	1489	1490	1490
1487	1491	1496	1499	1501	1502	1502	1503	1503
1508	1507	1506	1505	1505	1505	1505	1505	1505
2866	2913	2972	3003	3020	3028	3033	3037	3037
3116	3115	3116	3118	3120	3121	3122	3123	3123
3179	3177	3175	3173	3173	3172	3172	3171	3171
3205	3206	3207	3208	3208	3208	3209	3209	3209
3316	3317	3318	3319	3320	3320	3320	3321	3321

TABLE A.11: Electronic energies (E_h) for CBS extrapolation at selected points along the $H + C_2H_5$ reaction path with CASPT2/cc–pVQZ optimized geometries

S = 0.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–79.1185192	–79.4755693	–79.5083534	–79.5210325	–79.4810459	–79.4717598	–79.5124369
cc-pVQZ	–79.1238879	–79.5027729	–79.5312874	–79.5452879	–79.5084290	–79.4929874	–79.5355246
cc-pV5Z	–79.1252832						

S = 1.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–79.1219121	–79.4594153	–79.4910918	–79.5032501	–79.4639812	–79.4601319	–79.4965282
cc-pVQZ	–79.1272237	–79.4859862	–79.5134658	–79.5268838	–79.4907148	–79.4809998	–79.5191360
cc-pV5Z	–79.1286035						

S = 10.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–79.1229936	–79.4590165	–79.4900995	–79.5021620	–79.4639812	–79.4601319	–79.4965282
cc-pVQZ	–79.1283200	–79.4854302	–79.5122767	–79.5255766	–79.4907148	–79.4809998	–79.5191360
cc-pV5Z	–79.1297023						

TABLE A.12: Relative energies (ΔE, kcal mol^{-1}) at selected points along the H + C_2H_5 reaction path for each level of theory and basis sets leading to the CBS limit[a]

S = 0.0 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRPT2]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	+2.81	–10.39	–11.45	–11.84	–11.38	–7.53	–10.48
cc-pVQZ	+2.78	–10.88	–11.93	–12.37	–11.91	–7.84	–10.89
cc-pV5Z	+2.77	[–11.06]	[–12.10]	[–12.56]	[–12.09]	[–7.95]	[–11.04]
CBS	[+2.77]	[–11.24]	[–12.27]	[–12.75]	[–12.28]	[–8.06]	[–11.19]

S = 1.0 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRPT2]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	+0.68	–0.25	–0.62	–0.68	–0.67	–0.24	–0.50
cc-pVQZ	+0.69	–0.35	–0.75	–0.82	–0.79	–0.32	–0.61
cc-pV5Z	+0.69	[–0.39]	[–0.79]	[–0.87]	[–0.84]	[–0.35]	[–0.65]
CBS	[+0.69]	[–0.43]	[–0.84]	[–0.93]	[–0.88]	[–0.38]	[–0.70]

[a] The reference point is the energy of the separated reactants, taken from computations at S = 10.0 Å. CBS limits were evaluated using eqs 12 and 13 of the text.

A.2.5: $H + HCO \rightarrow H_2 + CO$

Table S13. Energy (kcal mol^{-1}) relative to separated reactants along the H + HCO reaction path for different levels of theory at the CBS limit[a]

S(Å)	CAS(2,2)	MRCI	MRCI+Q	CASPT2	Mk–MRPT2	Mk–MRCCSD	Mk–MRCCSD(T)
–0.2	–32.08	–39.79	–41.61	–37.85	–38.25	–42.40	–41.47
0.0	–22.53	–29.27	–30.89	–28.29	–29.46	–31.44	–31.36
0.2	–15.17	–20.46	–21.75	–20.21	–19.98	–22.70	–21.76
0.4	–9.81	–13.56	–14.48	–13.73	–13.66	–15.13	–14.83
0.6	–6.11	–8.54	–9.14	–8.86	–8.45	–9.71	–9.44
0.8	–3.70	–5.21	–5.58	–5.51	–5.00	–6.05	–5.83
1.0	–2.21	–3.14	–3.37	–3.37	–2.99	–3.66	–3.53
1.2	–1.32	–1.89	–2.04	–2.05	–1.76	–2.23	–2.16
1.4	–0.78	–1.14	–1.24	–1.25	–1.06	–1.35	–1.30
1.6	–0.46	–0.69	–0.76	–0.77	–0.63	–0.82	–0.80
1.8	–0.28	–0.42	–0.46	–0.47	–0.39	–0.48	–0.47
1.87	–0.23	–0.36	–0.39	–0.40	–0.32	–0.43	–0.42
1.95	–0.19	–0.29	–0.32	–0.33	–0.26	–0.35	–0.34
2.4	–0.06	–0.10	–0.12	–0.12	–0.09	–0.13	–0.13
2.7	–0.03	–0.06	–0.06	–0.07	–0.05	–0.07	–0.07
3.0	–0.01	–0.03	–0.04	–0.04	–0.03	–0.04	–0.04
5.0	0.00	0.00	0.00	0.00	0.00	0.00	0.09

[a]The geometries along the reaction path are optimized at the CASPT2/cc-pVQZ level of theory; the geometry at S = 10 Å is taken as the asymptotic limit.

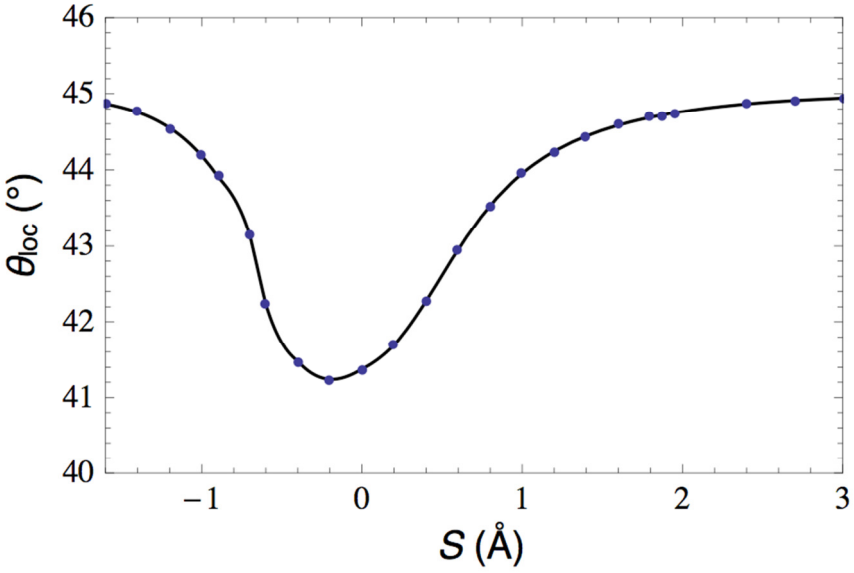


FIGURE A.6: Boys orbital localization angle θ_{loc} along the H + HCO reaction path.

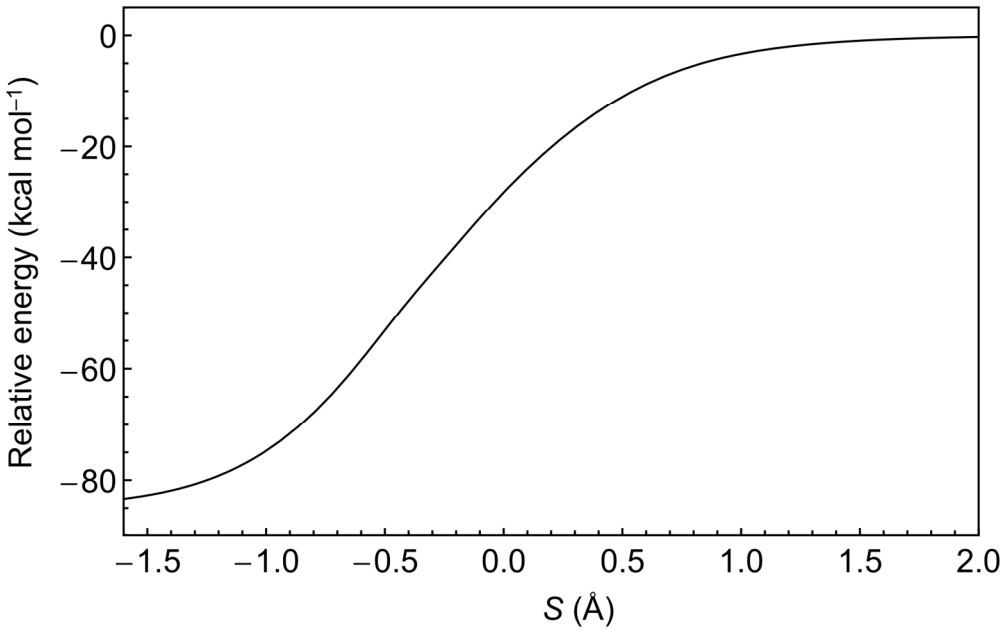


FIGURE A.7: The one-dimensional potential energy curve for H + HCO in CASPT2 level of theory.

TABLE A.14: Projected vibrational frequencies (ω, cm^{-1}) along the H + HCO reaction path

Reaction coordinate S(Å)								
–0.2	0.0	0.2	0.4	0.6	0.8	1	1.2	1.4
538	552	529	475	398	316	242	182	135
992	927	847	827	657	503	377	279	207
1242	1143	995	899	973	1013	1032	1043	1049
1499	1567	1795	1949	1925	1907	1897	1892	1889
2047	2020	1990	2057	2256	2422	2550	2638	2692
3159	3105	3077	3103	3104	3081	3049	3017	2986

Reaction coordinate S(Å)								
1.6	1.8	1.87	1.95	2.4	2.7	3	5	10
100	74	67	59	30	19	11	10i	9i
152	113	101	90	46	31	21	15	15
1053	1055	1056	1056	1058	1058	1059	1059	1059
1887	1886	1885	1885	1885	1884	1884	1884	1884
2724	2742	2746	2750	2760	2762	2763	2764	2764
2958	2935	2927	2920	2890	2877	2869	2861	2861

TABLE A.15: Electronic energies (E_h) for CBS extrapolation at selected points along the H + HCO reaction path with CASPT2/cc–pVQZ optimized geometries

S = 0.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–113.8223359	–114.2075754	–114.2168986	–114.2329475	–114.2130163	–114.1855074	–114.2228974
cc-pVQZ	–113.8310992	–114.2443057	–114.2489252	–114.2674881	–114.2492469	–114.2159838	–114.2555216
cc-pV5Z	–113.8330729						

S = 1.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–113.7895406	–114.1664907	–114.1721423	–114.1893712	–114.1734132	–114.1436736	–114.1789834
cc-pVQZ	–113.7984632	–114.2025019	–114.2043630	–114.2233808	–114.2095005	–114.1741813	–114.2115672
cc-pV5Z	–113.8005969						

S = 10.0 Å

	CAS(2,2)	Mk-MRPT2	Mk-MRCCSD	Mk-MRCCSD(T)	CASPT2	MRCISD	MRCISD+Q
cc-pVTZ	–113.7859422	–114.1620239	–114.1664876	–114.1839616	–114.1683280	–114.1388410	–114.1738660
cc-pVQZ	–113.7948820	–114.1978367	–114.1985824	–114.2178203	–114.2042254	–114.1692258	–114.2062744
cc-pV5Z	–113.7970492						

TABLE A.16: Relative energies (ΔE, kcal mol^{-1}) at selected points along the H + HCO reaction path for each level of theory and basis sets leading to the CBS limit[a]

S = 0.0 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRPT2]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	–22.84	–28.58	–31.63	–30.74	–28.04	–29.28	–30.77
cc-pVQZ	–22.73	–29.16	–31.59	–31.17	–28.25	–29.34	–30.90
cc-pV5Z	–22.61	[–29.28]	[–31.49]	[–31.24]	[–28.24]	[–29.28]	[–30.87]
CBS	[–22.53]	[–29.46]	[–31.44]	[–31.36]	[–28.29]	[–29.27]	[–30.89]

S = 1.0 Å

	ΔE[CAS(2,2)]	ΔE[Mk-MRPT2]	ΔE[Mk-MRCCSD]	ΔE[Mk-MRCCSD(T)]	ΔE[CASPT2]	ΔE[MRCI]	ΔE[MRCI+Q]
cc-pVTZ	–2.26	–2.80	–3.55	–3.39	–3.19	–3.03	–3.21
cc-pVQZ	–2.25	–2.93	–3.63	–3.49	–3.31	–3.11	–3.32
cc-pV5Z	–2.23	[–2.95]	[–3.64]	[–3.51]	[–3.34]	[–3.12]	[–3.34]
CBS	[–2.21]	[–2.99]	[–3.66]	[–3.53]	[–3.37]	[–3.14]	[–3.37]

[a] The reference point is the energy of the separated reactants, taken from computations at S = 10.0 Å. CBS limits were evaluated using eqs 12 and 13 of the text.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 4 [†]

[†] D. B. Magers, A. E. Vaughn, J. M. Turney, and W. D. Allen. To be submitted to *Journal of the American Chemical Society*.

A.1: Cartesian coordinates of 1-(1-diamantyl) diamantane at the MP2/cc-pVDZ level of theory listed in Å.

C	0.159422348	-0.328639304	-0.807321142
C	-1.079138505	-0.746897472	-1.687546509
C	-0.662961557	-1.000047660	-3.163263796
C	-0.139422513	0.306713404	-3.783358041
C	1.076131226	0.756822277	-2.946810180
C	0.637467876	1.012299375	-1.472314771
C	-0.453615107	2.108551157	-1.537282177
C	-1.671037206	1.620891890	-2.334371753
C	-2.190114486	0.321968311	-1.695355982
C	-1.246626368	1.364039893	-3.784413956
C	1.257677502	-1.409373266	-1.029325076
C	1.641697284	-1.628419666	-2.495454811
C	0.406258958	-2.089132626	-3.272834873
C	2.177585929	-0.303493128	-3.050698782
H	-1.487462496	-1.703568896	-1.322312425
H	-1.574796355	-1.317836726	-3.709227982
H	0.194149947	0.109314008	-4.822395302
H	1.452174956	1.720673155	-3.346153370
H	1.525624309	1.397797333	-0.939043851
H	-0.883351723	2.302750114	-4.244366572
H	-2.105446326	1.009948848	-4.386289288
H	-3.054748413	-0.061250183	-2.272454545
H	-2.554750352	0.528414900	-0.674070840
H	-0.784336546	2.434177561	-0.543755126
H	-0.018621131	2.999721271	-2.032242971
H	-2.463411073	2.392477289	-2.304528589
H	0.031142531	-3.042252861	-2.853021939
H	0.653118293	-2.270299391	-4.336440606
H	2.490066291	-0.416992486	-4.106448853
H	3.069626144	0.010854358	-2.475107712
H	2.174694285	-1.127468502	-0.486926865
H	0.908839743	-2.369474969	-0.603543419
H	2.430462547	-2.403111604	-2.542292507
C	-0.159422348	-0.328639304	0.807321142
C	1.079138505	-0.746897472	1.687546509
C	0.662961557	-1.000047660	3.163263796
C	0.139422513	0.306713404	3.783358041
C	-1.076131226	0.756822277	2.946810180

C	-0.637467876	1.012299375	1.472314771
C	2.190114486	0.321968311	1.695355982
C	1.671037206	1.620891890	2.334371753
C	0.453615107	2.108551157	1.537282177
C	1.246626368	1.364039893	3.784413956
C	-1.257677502	-1.409373266	1.029325076
C	-1.641697284	-1.628419666	2.495454811
C	-0.406258958	-2.089132626	3.272834873
C	-2.177585929	-0.303493128	3.050698782
H	1.487462496	-1.703568896	1.322312425
H	1.574796355	-1.317836726	3.709227982
H	-0.194149947	0.109314008	4.822395302
H	-1.452174956	1.720673155	3.346153370
H	-1.525624309	1.397797333	0.939043851
H	0.784336546	2.434177561	0.543755126
H	0.018621131	2.999721271	2.032242971
H	0.883351723	2.302750114	4.244366572
H	2.105446326	1.009948848	4.386289288
H	3.054748413	-0.061250183	2.272454545
H	2.554750352	0.528414900	0.674070840
H	2.463411073	2.392477289	2.304528589
H	-0.031142531	-3.042252861	2.853021939
H	-0.653118293	-2.270299391	4.336440606
H	-2.490066291	-0.416992486	4.106448853
H	-3.069626144	0.010854358	2.475107712
H	-2.174694285	-1.127468502	0.486926865
H	-0.908839743	-2.369474969	0.603543419
H	-2.430462547	-2.403111604	2.542292507