Electronic Structure of Excited States with Configuration Interaction Methods

by

Alan D. Chien

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2017

Doctoral Committee:

Assistant Professor Paul Zimmerman, Chair
Professor John Kieffer
Professor Roseanne J. Sension
Assistant Professor Dominika Kamila Zgid

Alan D. Chien

alandc@umich.edu

ORCID iD: 0000-0002-3103-0697

Dedication

To my parents, whose wisdom I benefit from every day.

Acknowledgements

During my time at the University of Michigan, I have encountered many great minds and made many great friends, all of whom have broadened my view of the world. First and foremost, I would like to express my sincere gratitude to my advisor Dr. Paul Zimmerman. Without his continuous support, guidance, and patience, I would not have been able to complete the work contained within this thesis. His hard work and quick mind are truly top notch, and I am forever thankful that he took the time to share his insights and critiques of my work. I would also like to thank my committee members for spending their time guiding yet another graduate student towards graduation, as well as their flexibility and kindness.

Of course, my lab mates also played a large role in the completion of my degree, not only offering technical help, but companionship as well. Special mention has to be made of Andy Vitek, a fellow Hoo and a wonderful friend to me for the past four years. From playing ping-pong to climbing down at the Red, spending time with him is always a nice refueling session that I look forward to. Other lab mates include: Mina Jafari, who literally refueled me, providing me with many snacks over the years. Though her habit of forcing me to clean my keyboard was a slight pain in the neck. Josh Kammeraad's analytical mind continues to astound me to this day. The clarity with which he sees problems is something I will (most likely) always remain envious of, as well as his never-ending cheerfulness and willingness to help. Laura Motta is a whirlwind of energy, whose enthusiasm for her work is infectious. Collaborating with her briefly exposed me to the world of relativistic quantum chemistry, which I may one day dive back into. Cody Aldaz helped me briefly fulfill the childhood dream of learning to skateboard, for which I am forever grateful. Ian Pendleton, Andrew Molina, and Jordan Metz were there from day one, pushing through the chaos that is graduate school with me, and a better crew to run with I could not have asked for. I would also like to thank all of the post-docs for their wide knowledge base that I often made use

of. Special mention goes to Yu Zhao, Jared Hansen, and Hyungjun Kim, who were good friends as well as sources of technical help. David Braun deserves a mention here as well, being the best System Administrator one could ask for.

There are many people outside of lab that I have spent ample time with and who deserve mention. The Crotchball crew from high school, whose friendship I will always treasure. Anthony Yeh and Stephen Cronk, my old college roommates and video game partners. Mitchell Smith, Matthew Wolf, and Tejas Navaratna, old and current roommates who made returning home a joy and not a chore. All of the staff at Planet Rock, whose hard work provides me with endless vertical terrain to scamper up. My numerous climbing buddies: Adam Neitzke, Josh Golec, Audrey Akcasu, Seth Elliot, Jenny Hebert, John Farr, Sin Sar Hsie, Jessie and Blake Tracy, Marissa Linne, Justin Li, and Laura Haskins to name a few, with whom I always enjoyed crushing at the crag or the gym. Special mention goes to Anita Luong, one of the few climbers that I've managed to convince to go lead climbing with me.

Finally, I would like to acknowledge my family, who have been there since the beginning. My dad, my mother, my brother, and my sister are all wonderfully supportive and inspiring people, whom I love dearly.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Appendices

Abstract

Computational chemistry is routinely applied to ground state molecular systems to provide chemical insights. Accurate excited state calculations, however, still typically require carefully tailored calculations and sizeable computational resources. This work focuses on the development of methods and strategies that enable the calculation of excited state properties with more accuracy and on larger systems than ever before. The first two Chapters focus on the spin-flip configuration interaction family of methods. Chapter 2 introduces us to the quantities one can obtain with excited state methods, with a challenging example being the electronic structure of a possible intramolecular singlet fission system, a quinoidal bithiophene. The study assigns an experimentally observed long-lived exciton to a long-lived singlet multiexciton state with a combination of energetic and transition dipole moment quantities. The spin-flip methodology is extended in Chapter 3 to provide more insight into the energetic orderings of the multiexciton states of a tetracene dimer, a model singlet fission system, showing that triplet decoupling should occur spontaneously upon population of the intermediate multiexciton state, ^{1}(TT). However, this extension enlarged the configuration spaces to the point that they became a limiting factor in the calculation of larger systems.

Therefore, the latter two Chapters focus on investigating new strategies for identifying and eliminating unneeded configurations. Chapter 4 presents iterative submatrix diagonalization, a procedure for converging the Davidson diagonalization procedure with a reduced set of active orbitals. This is accomplished by generating a systematic series of submatrix approximations to the full configuration space and solving for eigenpairs within the series until convergence of eigenpairs is achieved. The method shows promise, converging eigenvalues with a considerable reduction in orbitals and total computational time. Chapter 5 applies heat-bath configuration interaction towards obtaining exact excitation energies and examines various ways in which convergence is signified. A new convergence metric based on the magnitude of the perturbative correction is developed and converged excitation energies are obtained for systems as large as hexatriene. These results involved treating configuration spaces with as many as 10^{38} configurations, a full 29 orders of magnitude over what is achievable with conventional configuration interaction methods and 10 orders beyond results reported by other recent state-of-art solvers. While there is still a great deal of work to be done before excited state computational chemistry will be routinely applicable to a wide variety of systems, the various methods investigated and extended here show significant promise, especially those presented in the latter Chapters as these are generally applicable to any configuration interaction method.

Chapter 1: Introduction

1.1 Exploring Excited States with Computational Chemistry

The great appeal of computational chemistry lies in its promise of being able to accurately predict chemical properties of interest. To do so, quantum chemists endeavor to obtain state descriptions, or wave functions, that exactly solve the Schrödinger equation. However, computational chemistry is still far from being able to compute exact wave functions, from which the properties of a system are derived, in all but the smallest systems. This is due in large part to the correlated movement of electrons, where each electron's movement simultaneously affects the movement of all other electrons. Exactly accounting for this effect is difficult, limiting the routine application of exact treatments to diatomic systems. Nevertheless, quantum chemists have made considerable progress in generating approximate wave functions that are accurate enough for use in predicting and explaining chemical phenomena. A variety of electronic structure methods have been developed, mostly falling under the categories of configuration interaction (CI), perturbation theory (PT), coupled-cluster (CC), and density functional theory (DFT). This work expands the limits of CI methods, allowing computation of more accurate wave functions on larger systems, with an explicit focus on excited states.

In CI methods, wave functions are described using electron configurations, which are distributions of electrons in molecular orbitals. The application of CI methods to ground state phenomena can be routinely accomplished on a wide variety of systems, in large part because ground states are typically well described by a single electron configuration. This allows truncated CI methods to obtain qualitatively correct wave functions. Excited state calculations, however, are still a specialized endeavor, requiring tailored calculations, hefty computing resources, or both. This is because it is common for an excited state to be multi-reference in nature, requiring multiple configurations for a qualitatively correct description. Treating ground and excited states together with truncated CI methods then runs the risk of describing the states unequally, resulting in qualitatively incorrect excitation energies. A straightforward way to overcome different state natures is to obtain the exact wave function with an untruncated CI calculation, i.e. full configuration interaction (FCI). Unfortunately, this is not routinely possible because traditional FCI is very expensive, generally limited to systems about the size of a diatomic. Another, less computationally intensive strategy is to tailor calculations to the states of interest. Both strategies impede the routine application of excited state computational chemistry, requiring either specialized facilities or experienced quantum chemists. Method development in computational chemistry thus generally revolves around making methods less costly as well as more automated, reducing the need for manual setup.

This work focuses on exploring and expanding the capabilities of CI type methods, one of the earliest branches of electronic structure theory. Specifically, the first half of this dissertation involves the application and extension of spin-flip CI methods to the multiexciton states of singlet fission systems. The latter projects investigate optimizations of the algorithms that underlie all CI methods, pursuing strategies that any CI based method could use to become more efficient. We begin with a brief overview of current CI methodology and terminology in Section 1.2.

1.2 Theoretical Background

Here is provided a brief overview of the key equations and nomenclature that will most aid the reader in understanding the CI methodology employed in this dissertation. Much of the presented material is a summary of concepts spelled out in greater detail in References [1–3]. The interested reader is encouraged to seek out these texts if a deeper understanding is desired.

1.2.1 Schrödinger's Equation, Operators, and Wave functions

The aim of CI methods is to represent the wave function found in Schrödinger's equation (Eq. 1.1) as a linear combination of configurations. Schrödinger's equation defines a relation between the Hamiltonian operator, $\hat{H}$, and the wave function, $|\Psi\rangle$, stating that application of the Hamiltonian operator to the wave function returns the wave function times a constant, the energy of the system. This is exactly the form of an eigenvalue equation, for which practical strategies for finding the eigenvectors/values, i.e. wave functions and energies, are available from the field of linear algebra.

$$\hat{H}|\Psi\rangle = E|\Psi\rangle$$

Equation 1.1 Schrödinger equation

Put simply, an operator is a set of mathematical operations that have been grouped into a single symbol. The Hamiltonian operator is a set of operations that describes all interactions of a given system of atoms. The Hamiltonian operator (Eq. 1.2) is

$$\hat{H} = \sum_{A} \frac{1}{2M_A} \nabla_A^2 + \sum_{i} \frac{1}{2} \nabla_i^2 + \sum_{A>B} \frac{Z_A Z_B}{R_A - R_B} + \sum_{ij} \frac{1}{r_i - r_j} + \sum_{Ai} \frac{Z_A}{R_A - r_i}$$

Equation 1.2 Hamiltonian operator

where *i, j* refer to electrons and *A, B* to nuclei. The Hamiltonian includes five terms: nuclear kinetic energy, electronic kinetic energy, nuclear-nuclear repulsion, electron-electron repulsion, and nuclear-electron attraction. Typically, one simplifies this by taking the Born-Oppenheimer Approximation (BOA). The BOA treats the nuclei as stationary objects because the mass difference between nuclei and electrons causes the electrons to move at many times the speed of the nuclei. Thus, both purely nuclear terms in Equation 1.2 become constants and can be ignored. Equation 1.3 shows the BOA form of the Hamiltonian operator that is used in the quantum chemistry calculations described in this work.

$$\hat{H} = \sum_{i} \frac{1}{2} \nabla_i^2 + \sum_{ij} \frac{1}{r_i - r_j} + \sum_{Ai} \frac{Z_A}{R_A - r_i}$$

Equation 1.3 Born-Oppenheimer Approximation form of Hamiltonian operator

The wave function is a spatial function that describes the position of electrons in a system, whose square gives the electronic probability density. For single electron systems, such as a Hydrogen atom, wave functions have analytical solutions. Unfortunately, many-electron systems do not have analytical solutions since one must account for correlated electronic movements arising from the second term of the BOA Hamiltonian in Equation 1.3. CI methods treat correlation in an approximate manner by building up wave functions as linear combinations of electron configurations, which are distributions of electrons in molecular orbitals. Configurations are formally known as Slater determinants, which write electron configurations in a form that obeys the anti-symmetry rule required of fermionic (an electron is a type of fermion) wave functions. From here on, the terms "configuration" and "determinant" are used interchangeably. A central component of configurations are the molecular orbitals that electrons occupy, which are generated through a Hartree-Fock calculation.

1.2.2 Hartree-Fock, Basis Functions, and Molecular Orbitals

As a starting point for treating electron correlation, one may simplify the physical picture. Instead of having each electron interact simultaneously with all other electrons, the mean-field approximation assumes each electron interacts only with the mean electronic field generated by the other electrons in the system, leading to a single configuration picture of the wave function. The field is generated by the spatial locations of each electron, defined by a molecular orbital (MO). MO's themselves are composed of a linear combination of spatial functions termed basis functions provided by a given basis set. We can also take advantage of the variational principle, which guarantees that any approximate wave function's energy will be higher than the true energy. Thus, one can optimize MOs to give the best single configuration approximation to the true wave function by minimizing the single configuration's energy. This procedure is called Hartree-Fock (HF), and the optimization is accomplished with a self-consistent field (SCF) algorithm. HF calculations use basis sets with many more basis functions than are occupied by electrons to increase the flexibility of the MOs, allowing for more accurate single configuration approximations to the true wave function. HF is a key starting point for many higher-level methods, providing a set of one-particle basis functions (MOs) that can be used to expand the many-particle wave function. An illustration of the HF configuration showing electron occupations is shown in Figure 1.1.

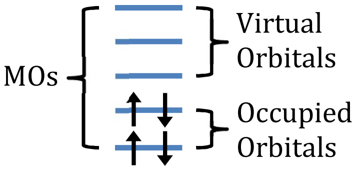


Figure 1.1 Illustration of HF determinant composed of five MOs (generated from five basis functions) and four electrons.

1.2.3 Configuration Interaction (CI) and Matrices in Quantum Chemistry

CI methods write the wave function as a linear combination of multiple configurations as in Equation 1.4 and illustrated in Figure 1.2. The extra configurations

$$|\Psi\rangle = c_{HF}\phi_{HF} + c_1\phi_1 + c_2\phi_2 + \cdots$$

Equation 1.4 Configuration interaction representation of a wave function

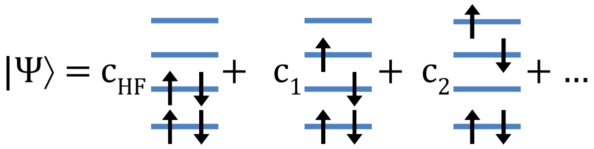


Figure 1.2 Pictorial depiction of a CI wave function composed of a linear combination of multiple determinants.

allow CI to explicitly take into account other possible electron positions, positions that were only treated in an average sense by HF. CI is a formally exact theory, meaning that if one had a complete basis set that spanned all space, and if one included all possible configurations in the wave function, one would have the exact wave function as a linear combination of configurations. Unfortunately, spanning all space is impossible given limited computational resources, and so we utilize finite basis sets, whose size is denoted as single-zeta (SZ), double-zeta (DZ), triple-zeta (TZ), etc. But even with finite basis sets, including all possible configurations in the CI space, i.e. Full Configuration Interaction (FCI), is very expensive. FCI scales exponentially with system size, limiting routine application to small systems such as diatomics. Alternative methods that produce FCI-quality energies at vastly reduced costs are an active area of research, and recent developments have allowed systems as large as hexatriene to be tackled, as will be discussed in Chapter 5. Of course, chemical space involves molecules many times larger than hexatriene, and so many truncated CI methods exist that limit their CI space by excitation level (CIS = CI + Single excitations, CISD = CI + Single + Double excitations, etc.) or other schemes in order to accomplish calculations on larger systems. However, truncation causes the loss of size-extensivity, in which absolute energies scale properly with system size. Figure 1.3 illustrates the hierarchy of CI methods as one approaches the true wave function.

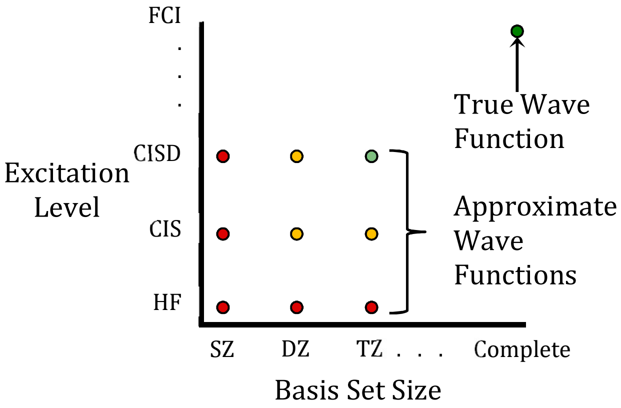


Figure 1.3 CI hierarchy of methods. CI methods become more exact as one raises the excitation levels included in the CI space and as one increases the basis set size, as indicated by the colors. The colors are included for illustrative purposes only, and should not be mistaken as indicating that HF is always a poor method of choice. In fact, HF methods give impressively accurate answers given the large approximations used in the method.

In CI methods, one solves the Schrödinger equation within a chosen CI space, which reduces to a matrix eigenvalue problem. This type of problem is well known in the field of linear algebra, and is solved by constructing and diagonalizing the Hamiltonian matrix. However, the straightforward diagonalization procedure becomes expensive, both in computer time and storage space, as one increases the dimension of the CI space - and Hamiltonians with dimensions greater than 10^8 are routinely encountered. The need for an efficient alternative diagonalization algorithm led to the development of the Davidson algorithm,[4–7] which takes advantage of the structure of the Hamiltonian – a sparse, symmetric, diagonally dominant matrix. The Davidson algorithm solves for the lowest few eigenvalues by iteratively expanding a subspace of the Hamiltonian until convergence is achieved. The Davidson algorithm, whose usefulness has made it ubiquitous throughout quantum chemistry methods, can still be improved upon with various strategies, as will be seen in Chapter 4.

1.2.4 CI for Excited States, FCI Approximations, and Spin-Flip Methodology

So far, we have neglected to explicitly discuss excited states in the context of HF and CI methods, as the focus has been on a basic understanding of the concepts and terminology. HF is a ground state method, focused on finding the MOs optimized for describing the ground state configuration that places all electrons in the lowest energy orbitals, also known as the HF configuration. CI's extra determinants both improve the description of the ground state and allow excited state wave functions to be obtained as higher roots in the matrix eigenvalue problem, with exact wave functions (within a given basis set) obtained with FCI. However, difficulties arise in getting accurate excitation energies with truncated CI methods due to the use of HF as the starting point, i.e. the use of an HF reference. The HF procedure only optimizes the occupied orbitals, meaning that the virtual orbitals occupied in excited states, are not explicitly optimized – they are merely byproducts of the optimization of the occupied orbitals. Thus, the MOs are biased towards providing a good description of the ground state. This leads to unbalanced treatments of the ground and excited states and poor excitation energies at lower levels of CI theory. This imbalance can be addressed in a number of ways.

The most straightforward approach, and the focus of Chapters 4 and 5, is to negate the effect of the unbalanced starting point by capturing all states to a high degree of accuracy with FCI-level wave functions. As traditional FCI is too expensive for routine application, methods that obtain FCI quality wave functions at a reduced cost have been extensively investigated. Some of the most promising developments are the incremental FCI (iFCI),[8–10] density matrix renormalization group (DMRG),[11–13] and FCI quantum Monte Carlo (FCIQMC) methods,[14–16] which take advantage of Hamiltonian structure and stochastic sampling to expand the system sizes that can be treated at the FCI level. DMRG[13] and FCIQMC[16] have been applied to butadiene in various basis sets, treating FCI spaces of

10^{25} and 10^{29} respectively. iFCI[10] is even more impressive, treating the FCI space of hexatriene, which contains 10^{38} configurations. This work investigates the capabilities of another method, heat-bath CI,[17,18] which is shown to be more flexible than any of these methods, obtaining accurate energy gaps for low-lying valence states for systems as large as hexatriene. However, there are many chemical systems of interest larger than hexatriene, requiring that other strategies for negating the unbalanced starting point of HF be utilized.

One such strategy is to utilize a different set of reference configurations. One may optimize MOs with respect to multiple configurations or choose a different reference configuration. Multi-configuration SCF (MCSCF) chooses to utilize multiple reference configurations. MCSCF is limited however by uncertainties in choosing the "correct" reference configurations and the increasing complexity of the orbital optimization procedure with more configurations. Another, simpler strategy is found in spin-flip (SF) CI methods, which utilize a single high-spin configuration as the reference.[19–24] The singly occupied frontier orbitals of high-spin references result in low-lying virtual orbitals being optimized for occupation. SF-CI methods therefore result in MOs that are well suited for multi-reference states which have significant contributions from occupation of frontier orbitals, such as in the description of bond-breaking (Figure 1.4) or certain low-lying excited states. The balanced reference allows accurate energy gaps to be obtained for certain states, allowing SF-CI methods to provide physical insights into systems much larger than hexatriene, such as a pentacene dimer.[25]

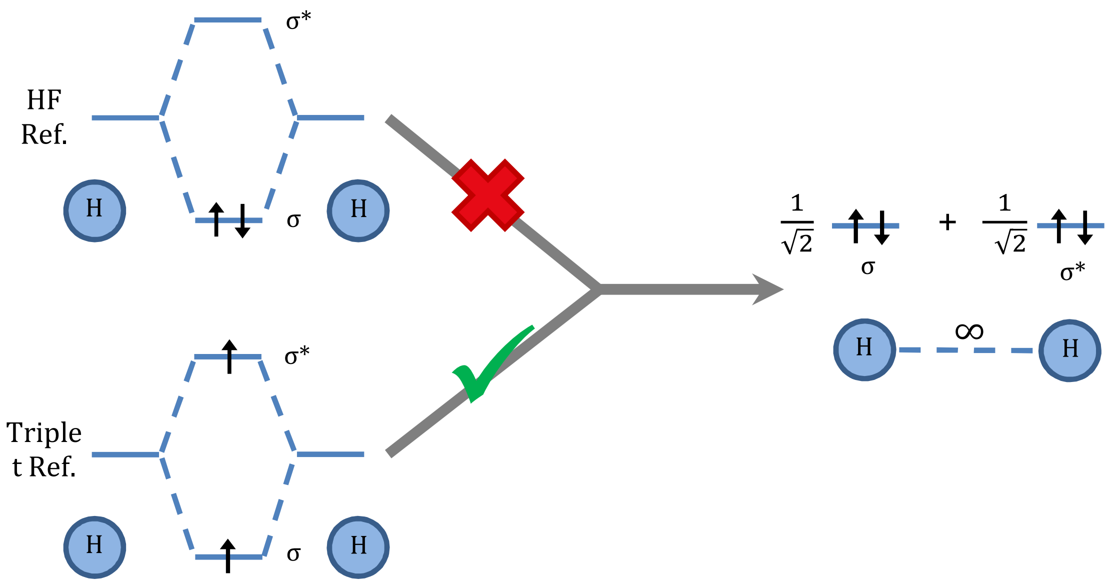


Figure 1.4 Depiction of the suitability of a high-spin reference for describing bond-breaking in an H_2 system. As the H_2 bond breaks homolytically, the frontier orbitals become degenerate, contributing equally to the ground-state wave function. An HF reference is imbalanced because it favors a single doubly occupied orbital. A triplet reference treats both orbitals equally, at all distances, resulting in a physically correct description when the bond breaks.

The main concepts and terminology of the CI methodology, and the challenges related to excited state calculations, have been reviewed. Following a brief outline of the rest of this dissertation, Chapters representing investigations into extending CI methodology towards more accurate results in ever larger CI spaces will be presented.

1.3 Dissertation Outline

In Chapter 1, a brief overview of computational chemistry was presented, with a focus on the difficulties one encounters in excited state calculations. A theoretical background of CI methodology was also presented as all of the methods improved upon in this dissertation are of the CI branch of electronic structure methods.

Chapter 2 presents a case study of the photodynamics and electronic structure of a quinoidal bithiophene using existing electronic structure methods.[26] The RAS(h,p)-SF method is used to characterize the electronic structure of a multiexciton dark state and the implementation and calculation of transition dipole moments provides evidence of a long lived singlet multiexciton state. This study provides a clear example of how computational studies can complement the results of experimental investigations.[27]

Chapter 3 presents an extension of the RAS(h,p)-SF method focused on increasing its accuracy for single-exciton states in an economical fashion. The extended method, RAS(2h,2p,S)-SF,[24] is applied towards a variety of systems with multi-reference states to quantify the effects of the extension. Most notably, the energetics of singlet, triplet, and quintet multiexciton states in a tetracene dimer are parsed and transition dipole moment calculations used to verify recent experimental results. This method is limited to systems about the size of tetracene dimers due to large CI spaces. As the size of the CI space is the bottleneck, strategies in eliminating insignificant determinants within CI spaces are explored in Chapters 4 and 5.

Chapter 4 introduces Iterative Submatrix Diagonalization (ISD), a method that cuts the cost of Davidson diagonalization by limiting the number of active orbitals used in generating the CI space. This scheme uses natural orbitals to order orbitals by estimated importance and then iteratively diagonalizes larger and larger orbital spaces until the wave functions are converged with respect to unfrozen orbitals. It is found that Davidson can converge wave functions with substantially smaller spaces and reduced timings. Furthermore, a well-behaved error control scheme is introduced to signal convergence.

Chapter 5 investigates the promise of another strategy for eliminating insignificant determinants with the recently introduced CI method, heat-bath CI (HCI). HCI is applied towards obtaining FCI-quality vertical excitation energies. HCI selects the important determinants from a large space based on their coupling to the primary determinants of a wave function, vastly reducing the determinants one must consider. HCI is applied to the FCI problem in the polyene series for systems up to hexatriene, with a FCI space of 10^{38}. A new metric is introduced and used to assess the accuracy of HCI energy gaps.

Chapter 6 includes Final Remarks, in which the findings of the prior Chapters are reviewed, and possible future projects laid out. There have been a great many improvements in CI methodology in the past couple of years. But the generation of novel strategies to cope with the large CI spaces that one routinely encounters in larger systems is most impactful, as these are strategies that can be implemented in almost all CI methodologies.

1.4 References

1. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Dover Publications, 1989).
2. Helgaker, T., Jorgensen, P. & Olsen, J. *Molecular Electron-Structure Theory*. (John Wiley & Sons, Inc., 2012).
3. Cramer, C. J. *Essentials of Computational Chemistry*. (John Wiley & Sons, Ltd, 2004).
4. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **17,** 87–94 (1975).
5. van Lenthe, J. H. & Pulay, P. A space-saving modification of Davidson's eigenvector algorithm. *J. Comput. Chem.* **11,** 1164–1168 (1990).
6. Murray, C. W., Racine, S. C. & Davidson, E. R. Improved algorithms for the lowest few eigenvalues and associated eigenvectors of large matrices. *J. Comput. Phys.* **103,** 382–389 (1992).
7. Leininger, M. L., Sherrill, C. D., Allen, W. D. & Schaefer, H. F. Systematic Study of Selected Diagonalization Methods for Configuration Interaction Matrices. *J. Comput. Chem.* **22,** 1574–1589 (2001).
8. Zimmerman, P. M. Incremental full configuration interaction. *J. Chem. Phys.* **146,** 104102 (2017).
9. Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *J. Chem. Phys.* **146,** 224104 (2017).
10. Zimmerman, P. M. Singlet–Triplet Gaps through Incremental Full Configuration Interaction. *J. Phys. Chem. A* **121,** 4712–4720 (2017).
11. White, S. R. & Martin, R. L. Ab initio quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* **110,** 4127–4130 (1999).
12. Chan, G. K.-L. & Sharma, S. The Density Matrix Renormalization Group in Quantum Chemistry. *Annu. Rev. Phys. Chem.* **62,** 465–481 (2011).
13. Olivares-Amaya, R. *et al.* The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* **142,** 34102 (2015).
14. Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131,** 54106 (2009).
15. Cleland, D., Booth, G. H. & Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **132,** (2010).
16. Daday, C., Smart, S., Booth, G. H., Alavi, A. & Filippi, C. Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question. *J. Chem. Theory Comput.* **8,** 4441–4451 (2012).
17. Holmes, A. A., Tubman, N. M. & Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **12,** 3674–3680 (2016).
18. Sharma, S., Holmes, A. A., Jeanmairet, G., Alavi, A. & Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **13,** 1595–1604 (2017).
19. Krylov, A. I. Size-consistent wave functions for bond-breaking: the equation-of-

motion spin-flip model. *Chem. Phys. Lett.* **338,** 375–384 (2001).

20. Krylov, A. I. Spin-flip configuration interaction: an electronic structure model that is both variational and size-consistent. *Chem. Phys. Lett.* **350,** 522–530 (2001).
21. Sears, J. S., Sherrill, C. D. & Krylov, A. I. A spin-complete version of the spin-flip approach to bond breaking: What is the impact of obtaining spin eigenfunctions? *J. Chem. Phys.* **118,** 9084–9094 (2003).
22. Casanova, D. & Head-Gordon, M. The spin-flip extended single excitation configuration interaction method. *J. Chem. Phys.* **129,** 1–13 (2008).
23. Zimmerman, P. M., Bell, F., Goldey, M., Bell, A. T. & Head-Gordon, M. Restricted active space spin-flip configuration interaction: Theory and examples for multiple spin flips with odd numbers of electrons. *J. Chem. Phys.* **137,** 164110 (2012).
24. Chien, A. D. & Zimmerman, P. M. Recovering dynamic correlation in spin flip configuration interaction through a difference dedicated approach. *J. Chem. Phys.* **146,** 14103 (2017).
25. Zimmerman, P. M., Bell, F., Casanova, D. & Head-Gordon, M. Mechanism for Singlet Fission in Pentacene and Tetracene: From Single Exciton to Two Triplets. *J. Am. Chem. Soc.* **133,** 19944–19952 (2011).
26. Chien, A. D. *et al.* Structure and Dynamics of the ^{1}(TT) State in a Quinoidal Bithiophene: Characterizing a Promising Intramolecular Singlet Fission Candidate. *J. Phys. Chem. C* **119,** 28258–28268 (2015).
27. Varnavski, O. *et al.* High Yield Ultrafast Intramolecular Singlet Exciton Fission in a Quinoidal Bithiophene. *J. Phys. Chem. Lett.* **6,** 1375–1384 (2015).

Chapter 2: Structure and Dynamics of the ^{1}(TT) State in a Quinoidal Bithiophene, Characterizing a Promising Intramolecular Singlet Fission Candidate

This Chapter largely based upon published work:
Reproduced with permission from A.D. Chien, A.R. Molina, N. Abeyasinghe, O.P. Varnavski, T. Goodson, and P.M. Zimmerman, *J. Phys. Chem. C* **119**, 28258 (2015).

2.1 Abstract

Singlet fission is a photophysical process of interest for increasing the efficiency of next-generation solar cells. Tetracyanoquinodimethane bithiophene (QOT2) has a long-lived (57 μs) photo-induced excited state that may correspond to triplets resulting from intramolecular singlet fission. Since singlet fission usually occurs through intermolecular processes, a detailed description of the excited states involved and their evolution is needed to verify this hypothesis. The photo response of QOT2 is investigated using high-level electronic structure methods and quantum dynamics simulations, which show ultrafast passage through a conical intersection from the bright 1^1B_u state to the dark 2^1A_g surface. Characterization of QOT2's 2^1A_g wave function found it to be composed of two strongly coupled triplets, leading to the first detailed electronic structure description of an intramolecular ^{1}(TT) state. The population of such a state upon excitation of QOT2 raises the possibility of singlet fission through conformational changes that decouple the triplets. However, reaching an appropriate geometry for decoupled triplets appears unlikely given an energy cost of 1.76 eV. Consequently, the hypothesis that the long-lived excited state corresponds to 2^1A_g, a spin singlet, strongly interacting double triplet was explored. Transition moment calculations to assign excited-state absorption signals and investigations into internal conversion and intersystem crossing decay pathways indicate

that a long-lived 2^1A_g with ^{1}(TT) character is consistent with the available experimental data.

2.2 Introduction

Understanding photophysical processes, such as singlet fission, that may be used in next-generation photovoltaic cells is one possible application of excited state computational chemistry that has significant relevance today, with burgeoning populations straining conventional energy sources. Singlet fission is a promising avenue for the improvement of photovoltaic cells as a means to surpass the conventional solar cell efficiency limit of 32%.[1] The singlet fission process downconverts high energy excitons by splitting the initial excited state into two lower energy excitons. By efficiently extracting these charge carriers, the maximum theoretical efficiency can be raised to 45% in cells which combine regular and singlet fission chromophores.[2,3] Some of the authors recently reported a study on tetracyanoquinodimethane bithiophene[4] (QOT2, Figure 2.1) that provided evidence towards intramolecular singlet fission, whereas most singlet fission processes are intermolecular.[5,6] Specifically, a long-lived (57 μs) exciton distinct from the native QOT2 triplet was observed along with a magnetic field dependence of emission. This system is of special interest due to the estimated 176% triplet yield,[4] which might be harnessed in next-generation solar cells.

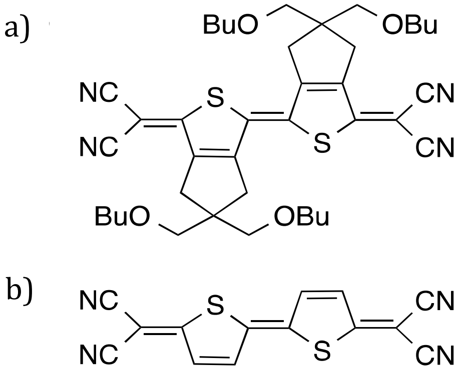


Figure 2.1 a) QOT2 chromophore. b) Model of QOT2 with β, β' substituents removed

The possibility of singlet fission in QOT2 has been suggested by ground and excited state characterization. First, the QOT2 ground state has biradicaloid character,[7–11] a property that helps satisfy the energy matching principle $2E(T_1) \leq E(S_1)$ for singlet fission.[12–14] studies of QOT2 excited states also suggest the existence of a dark state that can act as a singlet-coupled double triplet species, ^{1}(TT), populated from the initially excited bright state.[15,16] Kobayashi et al. conducted a transient absorption study on QOT2 and observed ultrafast passage (200 fs) to a low-lying dark state from the initially excited state.[15] A theoretical study on a related species, diphenoquinone bithiophene, showed multi-exciton character in its low-lying excited states.[16] This study then drew parallels between these states and polyenic doubly excited states, which have been described as two triplets coupled into an overall singlet,[17,18] similar to the ^{1}(TT) intermediate of singlet fission. These indications of QOT2 singlet fission were the driving force behind a recent QOT2 study which found evidence for intramolecular singlet fission.[4]

Much interest in singlet fission has focused on the mechanism for formation of the multiexcitonic ^{1}(TT) state from a bright singlet excited state. Analysis of model Hamiltonians led to the suggestion of indirect mechanisms that make use of an intermediate state characterized by charge-resonance configurations.[5,19] This mechanism is not immediately applicable to the intramolecular case as it involves coupling between at least two monomers. Furthermore, these model Hamiltonians include just four orbitals and can only provide a qualitative description of the states involved in singlet fission.[20] Quantitative descriptions require correlated many-electron wave functions for which one can develop further analyses.[21–23] One such method combines active space and spin-flip methodology, as reviewed in Chapter 1, Restricted Active Space Spin-Flip (RAS-SF).[24] RAS-SF is particularly well-suited for analyzing singlet fission systems, and has been used in the past to uncover an intermolecular singlet fission mechanism in which the ultrafast formation of a dark multi-exciton state is facilitated by vibronic coupling with the bright excited state.[25–28] This mechanism is especially relevant for QOT2, since passage through conical intersections may be able to explain both inter- and intra-molecular types of singlet fission.

In the previous study,[4] static electronic structure computations showed the initially excited 1^1B_u may relax to a lower-lying 2^1A_g state, consistent with the vibronic coupling mechanism. Further simulation indicated that 2^1A_g could transform via simultaneous rotation of the peripheral dicyanomethylene groups, but independent triplet formation would have a high cost of ~1.6 eV. Nonetheless, it was suggested that an unknown, lower cost path to an uncorrelated triplet pair involving dicyanomethylene distortions may exist – thereby explaining the experimentally observed long-lived spectroscopic signals.[4]

Further investigations are required to substantiate how and to what extent intramolecular singlet fission occurs in QOT2. High-level computational methods are herein used to provide insight into QOT2 photo physics by characterizing electronic states and ultrafast wave packet dynamics. Extended multi-state CAS second-order perturbation theory (XMS-CASPT2) calculations suggest an excited-state relaxation pathway through the conical intersection region between the bright 1^1B_u state and dark multi-excitonic 2^1A_g state. This pathway was shown to be viable using exact quantum dynamics in a four-mode, linear vibronic coupling Hamiltonian model (coupled to an Ohmic bath) that was created directly from *ab initio* calculations. The possibility of singlet fission was then investigated by characterizing the 2^1A_g state. Decomposition of 2^1A_g using a CAS(4,4) wave function reveals it to be composed of two strongly coupled triplets which are incompletely localized to either half of QOT2. Being composed of two triplets, QOT2's 2^1A_g bears similarities to the commonly invoked intermolecular ^{1}(TT) state in singlet fission systems, but it will be shown below that it differs by having strong electronic coupling between the triplets. Other possible decay channels for the 2^1A_g state, such as internal conversion to the ground state and intersystem crossing to triplet states, were evaluated. These investigations indicate that the rapidly generated dark 2^1A_g exciton may be long lived due to inhibition of internal conversion and intersystem crossing by the rigid QOT2 geometric structure. Finally, *ab initio* calculations provide assignment of experimental absorption signals to the 2^1A_g state and suggest that complete electronic decoupling of the ^{1}(TT) triplets within a QOT2 monomer is likely infeasible.

2.3 Methods

2.3.1 Computational Details

Unless otherwise mentioned, all calculations were performed on the model molecule (Figure 2.1b) which omits the β, β' substituents on the thiophene rings. These bis(butoxymethyl)cyclopentene substituents were initially introduced for their solubilizing properties.[29, 30] The neutral character of these moieties and their lack of conjugation with the π-backbone suggests that they play a negligible part in the low-lying π-π* states of interest. This view is supported by density functional theory studies which showed that inclusion of the β, β' substituents had a negligible effect on ground state geometries and HOMO/LUMOs in various members of the QOTn series.[7,9] The model molecule retains the C_{2h} symmetry of QOT2.

All geometries were optimized using the XMS-CASPT2[31] method with a 6-31G* basis set as implemented in Molpro.[32] XMS-CASPT2 provides more accurate energetics than standard CASPT2 near conical intersections or avoided crossings.[15] C_{2h} symmetry was employed and an active space of 10 electrons in 8 π orbitals was used. For all calculation parameters and rationalization of active space, see section A.1.1 of Appendix A.

The ultrafast relaxation dynamics between the 1^1B_u and 2^1A_g states were investigated using the Multi-Configurational Time-Dependent Hartree (MCTDH) package.[33,34,35] The wave packet is propagated on a two-state four-dimensional linear vibronic coupling Hamiltonian (LVCH) with a weakly coupled heat bath of 20 harmonic oscillators. Three of the modes were symmetric modes that tuned the energy gap. The fourth, asymmetric, mode introduced coupling between the two states of interest by lowering the symmetry of QOT2. To obtain the LVCH parameters, one-dimensional scans along all four vibrational modes were generated using the RAS-SF method[24] with 6 electrons in 6 π-orbitals, a septet high spin reference, and the 6-31G* basis. The parameters were then obtained by minimizing the residual squared error via gradient descent. An Ohmic bath of harmonic oscillators is used to model energy dissipation from interaction with the environment. Dynamics calculation details are found in Appendix A, section A.1.2

Transition dipole moment (TDM) calculations were used to assign experimental absorption signals to specific state transitions. XMS-CASPT2 TDMs are not implemented in MOLPRO and CASSCF TDMs showed considerable deviations with minor active space variations (Figure A.2), so neither method could be reliably used to assign absorption strengths. CASSCF TDMs are likely unreliable because simulations have shown that inclusion of σ-π dynamic correlation increases the accuracy of TDM calculations in polyenes,[36] which would require a large active space. Given the polyenic character of quinoidal oligothiophenes[16] and the larger amount of dynamic correlation in RAS-SF versus CASSCF, RAS-SF TDMs were implemented in a development version of the Q-Chem 4.0[37] software package for this study and found to be much more consistent than CASSCF TDMs.

For complete details on the other computations discussed below (spin-orbit coupling, localized orbitals, constrained DFT, and classical dynamics), see Appendix A and the Supporting Information (SI) of published work.[38]

2.3.2 Experimental Details

The computational investigations were supported by a variety of spectroscopic data.[4] A microsecond transient absorption spectrum of QOT2 using a flash photolysis setup detected the long-lived (57 μs) species of interest. A two-color pump-probe (445nm pump, 890nm probe) transmission experiment, in which attenuation of the probe beam was measured at various pump powers, demonstrated strong absorption to low-lying excited states. The effect of magnetic fields on the time-resolved fluorescence signal was also measured and shown to be significant. Finally, a time-correlated single photon counting setup was used to detect delayed fluorescence of an excited QOT2 sample. Additional synthetic and experimental details can be found in previously published works.[4,38]

2.4 Results and Discussion

2.4.1 Immediate Evolution of the Bright 1^1B_u Exciton

The first step in investigating the photo dynamics of QOT2 is to determine the fate of the initially excited bright 1^1B_u exciton. Previously reported femtosecond pump-probe experiments show complete depopulation of 1^1B_u to a longer-lived dark state within 1 ps.[4] Kobayashi et al. used a high time-resolution experiment to assign a decay time of 200 fs to the same transition.[15] The ultrafast nature of 1^1B_u depopulation suggests passage through a conical intersection, although the dark state remains unidentifiable by these data alone. To help identify possible conical intersections, the states neighboring 1^1B_u need to be identified.

For the key low-lying states of QOT2, XMS-CASPT2 was used to map the surrounding energetic landscape (Figure 2.2). XMS-CASPT2 is particularly well-suited for this system because it is more accurate than standard CASPT2 near conical intersections, and both methods are able to capture single- and multi-excitonic states.[31] These results allow us to determine that the accessible states following photo excitation of QOT2 are 1^1A_g, 2^1A_g, 1^1B_u and 1^3B_u. The photo response of QOT2 begins with excitation of the ground state, 1^1A_g, which leads to population of the optically allowed bright singlet, 1^1B_u.

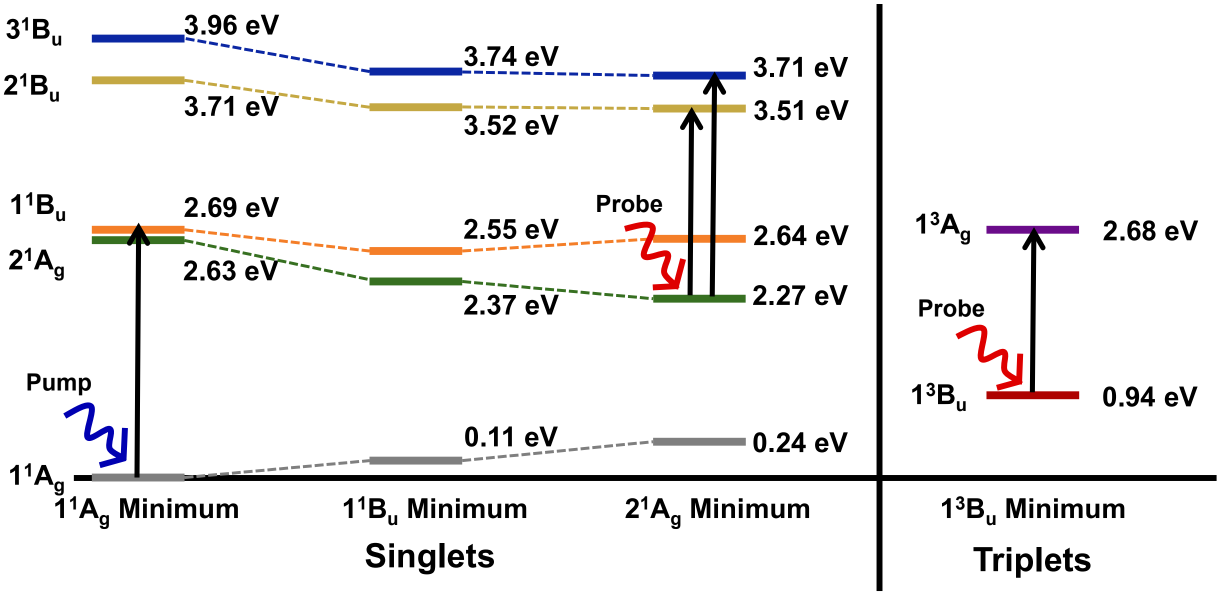


Figure 2.2 Energy landscape from XMS-CASPT2 (10,8) geometries and corresponding vertical excitation energies. Important features to note are the near degeneracy between 1^1B_u and 2^1A_g at the 1^1A_g geometry and the increasing energy gaps at the 1^1B_u and 2^1A_g geometries. Arrows indicate possible light-driven transitions.

At the vertical excitation geometry, the proximity of the dark 2^1A_g state suggests that 1^1B_u's evolution is critically influenced by a nearby conical intersection, which will be discussed in more detail below. The calculated n^1B_u and mono-triplet states are not immediately accessible but are useful for assigning the excited state absorption signals that will be discussed later. A vertical energy gap of 2.69 eV was found for the 1^1A_g to 1^1B_u excitation. This gap correlates well with the band maximum of 2.25 eV obtained from UV-vis spectra.[4] In general, band maxima are not expected to be equivalent to vertical excitations, but the proximity of the two values indicates a good description of the initial photo excitation of QOT2. The adiabatic energy gap between 1^1A_g and T_1, 1^3B_u, is 0.94 eV, in agreement with singlet oxygen sensitization experiments that place an upper bound of 0.98 eV on the triplet energy.[4] In addition to the agreement with experimental measurement, these values are qualitatively similar to existing electronic structure calculations on QOT2.[4]

Figure 2.2 implies a possible pathway for depopulation of the bright 1^1B_u state. The separation of only 0.06 eV between 1^1B_u and 2^1A_g upon vertical excitation suggests the presence of a nearby conical intersection.[4,15,25–27] Indeed, RAS-SF scans (Figure A1) show that the states cross along specific symmetric modes. At 1^1B_u and 2^1A_g relaxed geometries, the 2^1A_g state is considerably lower in energy than the 1^1B_u state. This indicates that population transfer may be rapidly downhill and that return to the 1^1B_u surface will be less favorable after transition to the dark 2^1A_g surface.

The role of the conical intersection in ultrafast conversion from the bright 1^1B_u to a dark state was examined with quantum dynamics simulations using MCTDH.[33,34,35] A two-electronic-state LVC Hamiltonian (see Computational Details) incorporating one asymmetric coupling mode and three symmetric tuning modes (Figure 2.3) was employed to capture the key electronic and vibrational degrees of freedom in the system. The inclusion of an Ohmic bath accounted for dissipative interactions with the solvent. The tuning modes adjust the energy gap between the two states and are Condon-active, symmetric vibrations that preserve C_{2h} symmetry. The asymmetric coupling mode, chosen because it coincides with the non-adiabatic coupling vector as calculated at the CASSCF(4,4) level, captures the coupling between 1^1B_u and 2^1A_g.

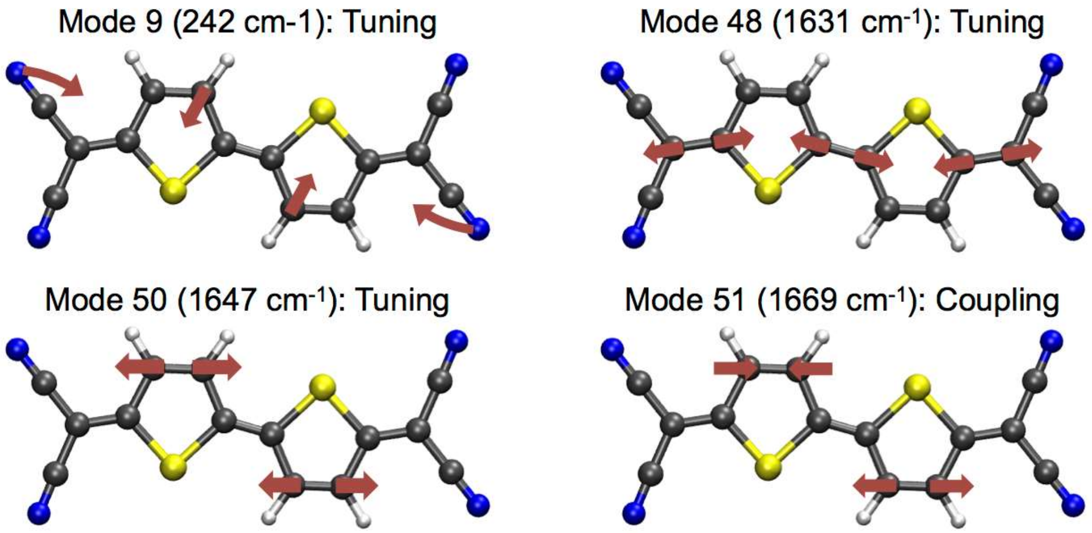


Figure 2.3 Major movements of the four modes involved in the initial photo dynamics. Note the asymmetry present in the coupling mode, breaking symmetry and allowing coupling between the 2^1A_g and 1^1B_u surfaces.

RAS-SF(6,6) scans along the tuning modes span the symmetry-allowed conical intersection region (Figure A1). The asymmetric coupling mode introduces interaction between the two electronic states and drives population transfer. Motion along this mode is critical because there is no coupling between the 1^1B_u and 2^1A_g until asymmetric vibrational motion breaks the electronic state symmetry. Overall, MCTDH allows exact dynamics simulations within the given LVC Hamiltonian model, giving a meaningful quantum mechanical picture for this ultrafast process.

The quantum dynamics simulations (Figure 2.4) indicate rapid population transfer to the dark 2^1A_g state with a 76 fs decay time, in reasonable agreement with the 200 fs decay time obtained by Kobayashi et al.[15] Population transfer is largely driven by motion in modes 48 and 50, which can be understood by the shape of their potential energy surfaces (Figure A.1). Starting on the 1^1B_u surface, motion along these modes passes through the conical intersection, resulting in the accumulation of kinetic energy, 0.45 eV and 0.7 eV for modes 48 and 50, respectively. This corresponds to roughly six additional vibrational quanta in modes 48 and 50 after the initial transfer to the dark 2^1A_g (Figure A.3). The extra energy may allow QOT2 to re-access the conical intersection for a period of time before energy dissipation occurs through bath interactions. This simulation establishes that ultrafast depopulation of the bright 1^1B_u exciton[4,15] occurs due to passage through a conical intersection to the dark 2^1A_g surface. Having demonstrated that 2^1A_g is populated within the first picosecond after excitation, we now examine properties of the 2^1A_g excited state at later timescales.

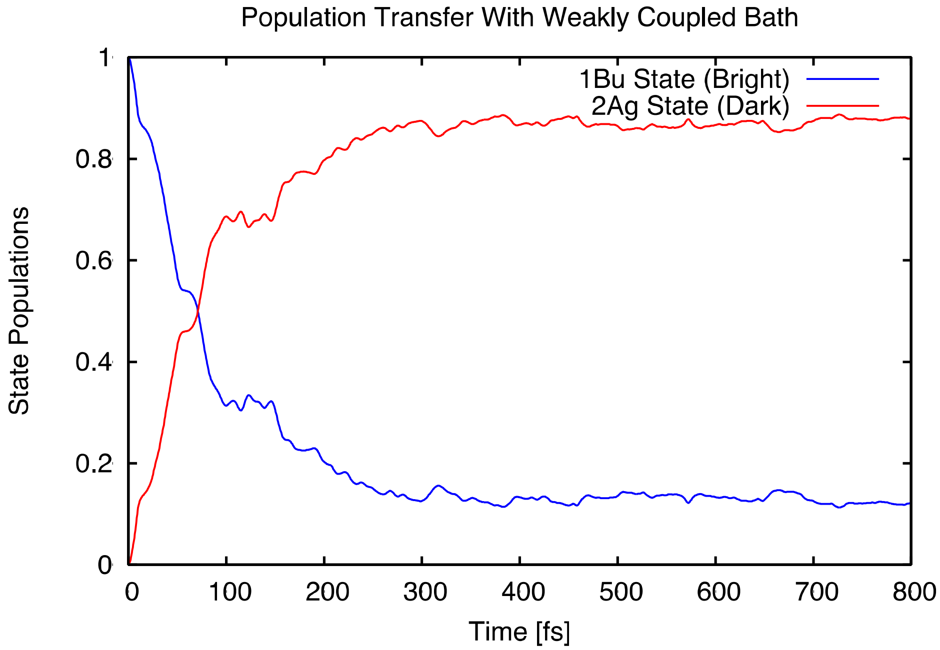


Figure 2.4 Two electronic state quantum dynamics using four vibrational modes, three tuning and one coupling, to model population transfer between the 1^1B_u and 2^1A_g surfaces in the presence of an Ohmic bath.

2.4.2 Nature and Transformations of the 2^1A_g Excited State

Microsecond transient absorption spectra probing excited state transitions to high-lying states show a signal corresponding to a photo induced state with a 57 μs lifetime.[4] Since population transfer from 1^1B_u to 2^1A_g occurs on the order of femtoseconds, this long-lived state is likely evolved from the 2^1A_g state. Unfortunately, XMS-CASPT2 assignments of high-lying QOT2 states are difficult due to the need to account for many states in the calculations. Thus, for guidance on the evolution of 2^1A_g, we examine a nanosecond pump-probe (445 nm pump, 890 nm probe) transmission experiment that probes excited state transitions to low-lying states.[38] Simulation of these low-lying transitions involves fewer states and thus can be readily accomplished by XMS-CASPT2 calculations.

Absorption of the low-energy probe (890 nm, 1.39 eV) is energetically consistent with transitions corresponding to two distinct possibilities. The first option, suggested by the observation of an emission increase when a magnetic field is applied[38] is the formation of a triplet-like species. The $T_1 \rightarrow T_2$ transition at 1.74 eV (Figure 2.2) is near the probe energy of 1.39 eV and in principle could be the source of this signal. Alternatively, the $2^1A_g \rightarrow n^1B_u$ transitions, which range from 0.37 eV to 1.44 eV, could also be the source of probe absorption, indicating a persistent, long-lived 2^1A_g exciton. However, excited singlet lifetimes are usually far shorter than the observed microsecond lifetime and there is yet no explanation for the observed magnetic dependence on emission. Previous investigations that observed a lack of solvent polarity and viscosity dependence indicates that the detected species is not the result of intermolecular processes or a charge-transfer species.[4] Thus, intersystem crossing from 2^1A_g to the triplet manifold is first considered, where there exists precedent for microsecond lifetimes and magnetic field effects.

2.4.2.1 Intersystem Crossing from 2^1A_g

The sulfur in oligothiophenes can introduce significant spin-orbit coupling (SOC) and increased efficiency of intersystem crossing (ISC).[39–41] Studies on ISC in oligothiophenes[42–44] and other aromatic systems[45] have found that non-planar conformations contribute the most to ISC. In QOT2, however, non-planar conformations will be diminished due to its quinoidal nature, which increases the rigidity of the conjugated backbone.

QOT2 ISC is most likely to occur from 2^1A_g since the bright 1^1B_u quickly converts to 2^1A_g by passing through a conical intersection, as demonstrated above. To identify possible ISC products, both SOC and energy gaps must be considered. Given that ISC is starting from 2^1A_g, a π-π* singlet, the El-Sayed rules[46] predict that σ-π* triplets will have significant SOC's. To locate low-lying σ-π* triplets, TDDFT calculations at the 2^1A_g geometry were run since the XMS-CASPT2 calculations only targeted π-π* states. However, it was found that all low-lying triplets are of π-π* character, with the nearest σ-π* triplet located 2 eV above 1^1B_u. This large energy gap indicates that σ-π* states will play an insignificant role in ISC from 2^1A_g. Possible products of QOT2 ISC are thus restricted to the energetically nearby π-π* triplets, 1^3B_u and 1^3A_g.

Transitions from 2^1A_g to 1^3B_u are forbidden by symmetry selection rules,[42,47,48] suggesting that the direct crossing to 1^3B_u is unlikely. For the 2^1A_g to 1^3A_g transition, the SOC was computed to be 0.0043 cm^{-1} at the planar 2^1A_g geometry. The relaxed 1^3A_g lies 0.14 eV *above* 2^1A_g, so minimal ISC between these states is expected. A Marcus[49] theory prediction of the rate gives an ISC time constant of 18.9 ms (Appendix A.2) The smallest time constant possible at this value of SOC—when the reorganization energy exactly cancels out the energy change—would be 81.2 μs. These values indicate that ISC from 2^1A_g will play an insignificant role in QOT2 dynamics from the planar 2^1A_g geometry.

Although non-planar conformations are expected to play a minimal role due to the rigid quinoidal backbone, central C=C bond twists will increase SOC values compared to the planar 2^1A_g geometry. These increases, however, were not large enough to yield significant changes in ISC rates (Table A.2). Furthermore, the estimated 176% triplet yield in QOT2[4] suggests ISC to be inoperative since it can generate only one triplet per excited QOT2.

2.4.2.2 Intramolecular Double Triplet, $^1(TT)$, States and 2^1A_g

To date, the rationale behind efficient intramolecular singlet fission systems has been centered around structures which allow the $^1(TT)$ state to spatially separate, and thus electronically decouple, into two independent triplets. The first design of covalently tethered tetracenes generated triplet yields of only around 3%,[50] but recent efforts show higher yields. For instance, triplet quantum yields of 156% in covalently tethered pentacene dimers[51] and 173% in a benzodithiophene – thiophene-1,1-dioxide block copolymer[52] were recently observed. The above systems differ from QOT2 in that spatial separation of the two triplets generated by singlet fission is highly plausible. Specifically, the pentacene dimer can localize one triplet to each pentacene fragment, and the block copolymer can separate triplets along the polymer chain. While there are reports of carotenoids undergoing intramolecular singlet fission when bound to light-harvesting complexes,[53,54] singlet fission does not occur as separated monomers in solution.[53] The structure of QOT2 is not amenable to similar mechanisms of triplet separation due to strong coupling along the conjugated backbone. It is thus possible that singlet fission in QOT2 may occur through a novel intramolecular mechanism.

The energy matching condition for singlet fission, $E(S_1) \geq 2E(T_1)$,[5,6] is fulfilled in QOT2. Experimentally, $E(S_1) = 2.25$eV and $E(T_1) < 0.98$eV from UV-Vis spectra and oxygen singlet sensitization, respectively,[4] satisfy the condition. XMS-CASPT2 energies (Figure 2.2) indicate that 1^1B_u and 2^1A_g states also fulfill the condition. This is consistent with previous studies that noted the biradicaloid character of the QOTn series[7–11] and predicted the ability of biradicaloid species to satisfy the energy matching criteria.[12–14]

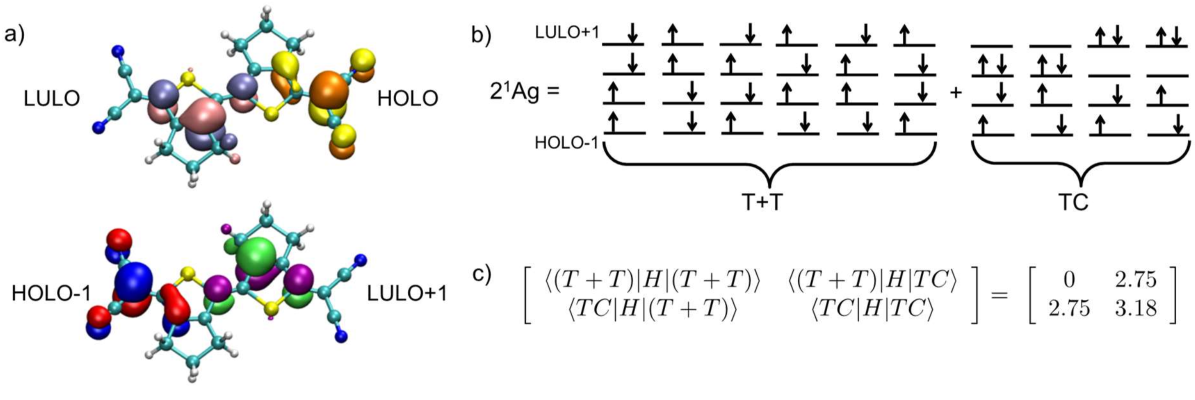
Multiexcitonic ^{1}(TT) states, which are double triplets spin-recoupled into an overall singlet, can have a range of electronic coupling strengths between the two triplets. In intermolecular singlet fission systems, these ^{1}(TT) states eventually evolve into spatially separated, electronically decoupled, independent triplets – at which point the ^{1}(TT) state becomes degenerate with a ^{5}(TT) state. Studies of pentacene show that vibronic coupling can generate these multi-excitonic ^{1}(TT) states.[25–27] Where a quintet (i.e. ^{5}(TT)) has four singly occupied orbitals, ^{1}(TT) is expected to have slightly changed occupancies when the triplets are proximate to one another. Correlations among the four electrons are responsible for this effect, so the singly occupied limit will be observed only in spatially separated singlet-coupled triplets. To illustrate the electronic coupling in ^{1}(TT) states, RAS-SF natural orbital (NO) occupancies of a quintet are compared to the dark ^{1}(TT) state of a pentacene dimer embedded in the crystal environment[26] (Appendix A.3) and 2^1A_g of QOT2 in Table 2.1. In the embedded pentacene dimer and QOT2's 2^1A_g state, the deviation from the unit occupation of orbitals seen in a quintet indicates the presence of electronic coupling between the two triplets.

Table 2.1 Natural orbital occupancies for various multi-exciton systems demonstrating the deviation from quintet single occupations.

	QOT2[a] Quintet	QOT2[a] 2^1A_g	Pentacene[b] Dimer ^{5}(TT)	Pentacene[b] Dimer ^{1}(TT)
LUNO+1	1.00	0.34	1.00	0.74
LUNO	1.00	1.11	1.00	0.99
HONO	1.00	0.95	1.00	1.01
HONO-1	1.00	1.61	1.00	1.26

[a] Obtained with RAS-SF(10,8).
[b] Obtained with RAS-SF(4,4).

In order to quantify the strength of triplet coupling in QOT2's 2^1A_g state, localized orbitals (LOs) were created. Specifically, four key frontier orbitals (Appendix A.4), HOMO-1 through LUMO+1, were localized to obtain their LO analogues, HOLO-1 through LULO+1, for use in a Spin-Flip CAS-CI(4,4) calculation. Figure 2.5 shows that the LOs are not fully localized to either half of QOT2, i.e. the four singly occupied orbitals cannot be fully spatially disentangled. This indicates the inability of the triplets to completely decouple from one another at the 2^1A_g geometry, as the spatial overlap allows for direct coupling through exchange interactions. These LOs allow diabatization of 2^1A_g into independent, decoupled triplet (T+T) and triplet coupling (TC) contributions. In this local orbital basis, the quintet state has four orbital occupancies of 1.00, meaning that the (T+T) component represents 100% of the wave function. In 2^1A_g these configurations make up 74.6% of the wave function, indicating a large independent triplet character (Figure 2.5). The remaining electronic configurations make up the TC diabat, which is dominated (77% of its 25.4% fraction) by four configurations involving interaction between the spatially overlapping LULO and LULO+1 orbitals. TC therefore describes the most significant electron correlations that are responsible for the non-unit occupations of Table 2.1. The diabatic Hamiltonian reveals that the off-diagonals between the T+T and TC components are of the same magnitude as the energy difference between the two (Figure 2.5c). This mixing allows us to determine that QOT2's 2^1A_g is a ^{1}(TT) state with significant electronic coupling present between the triplets, $\Psi_{2Ag} \approx 0.864\,(T+T) + 0.504\,(TC)$.



$$\begin{bmatrix} \langle (T+T)|H|(T+T)\rangle & \langle (T+T)|H|TC\rangle \\ \langle TC|H|(T+T)\rangle & \langle TC|H|TC\rangle \end{bmatrix} = \begin{bmatrix} 0 & 2.75 \\ 2.75 & 3.18 \end{bmatrix}$$

Figure 2.5 a) Localized frontier orbitals in QOT2. b) Decomposition of 2^1A_g into (T+T) and TC components in the localized orbital basis in a (4,4) active space. The spatial orbitals are HOLO-1, HOLO, LULO, LULO+1 from top to bottom in the determinants. c) Diabatic Hamiltonian of the 2^1A_g wave function demonstrating the strong mixing between two configurations. Values shown are in eV.

The ^{1}(TT) state described above is not the expected final form of triplets in singlet fission, which is often considered complete when the coupled triplets in ^{1}(TT) have evolved into spatially separated, electronically decoupled triplets. Although spatially separated, the triplets can be spin-entangled and interact with one another to create observable phenomena such as quantum beats in delayed fluorescence spectra.[55] These beats are difficult to observe even in known singlet fission systems, so the fact that no such beats are observed in QOT2's delayed fluorescence does not directly indicate a lack of singlet fission. Decoupling the triplets in the 2^1A_g wave function would require complete spatial separation of the LOs. This could be accomplished by creating perpendicular π-systems through rotation of QOT2's C-C bonds. To determine if 2^1A_g can evolve from strongly coupled to decoupled triplets on a single QOT2, we searched for QOT2 geometries consistent with two independent triplets.

To obtain a decoupled triplet structure, a constrained Hartree-Fock geometry optimization was performed in which an α triplet was localized to one half of the model molecule and a β triplet to the other half. The resulting geometry exhibited separated π-systems resulting from rotation of both cyano groups to positions perpendicular to the thiophene rings (Figure 2.6a). RAS-SF calculations at this twisted geometry (Figure 2.6b) lead to 2^1A_g natural orbitals with four singly occupied orbitals and a near-degeneracy between 2^1A_g and 1^5A_g, indicating that 2^1A_g becomes composed of independent triplets. Travelling to this decoupled triplet geometry along the 2^1A_g surface has an energetic cost of 1.76 eV, similar to that found in a previous investigation (1.6 eV).[4] Thus, evolution of the 2^1A_g into two independent triplets through a simultaneous twist of the terminal dicyanomethylene groups is an activated process with a high barrier, and therefore unlikely.

Further evidence against triplet decoupling comes from the difference between the twisted, decoupled triplet geometry and the planar, mono-triplet geometry. These two geometries have different π orbital structures, indicating that the separated triplet $T_1 \rightarrow T_2$ transition will most likely have different energetics compared to the mono-triplet $T_1 \rightarrow T_2$. However, the transient absorption maximum lies close to that of the triplet-sensitized absorption spectra, being red-shifted by only 30nm.[4] This coincident position suggests that a twisted decoupled triplet geometry is not responsible for the signal, indicating that this geometry is not reached in QOT2.

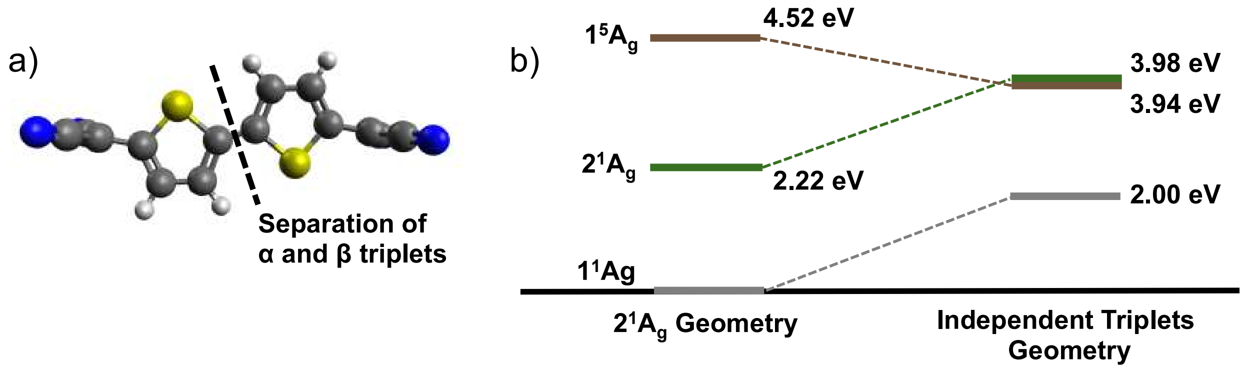


Figure 2.6 a) Electronically decoupled ^{1}(TT) structure with two cyano groups out of plane by 90 degrees. b) Energies resulting from RAS-SF(10,8) calculations at the 2^1A_g minimum and the independent triplet geometry.

2.4.3 The Possibility of Long-Lived 2^1A_g

Having obtained evidence against ISC and geometric distortions leading to a ^{1}(TT) state with electronically decoupled triplets, the intriguing possibility remains that excited state absorptions seen in the microsecond transient absorption spectra[38] arise from transitions from a long-lived 2^1A_g singlet state. This viewpoint would need to be consistent with the strong nanosecond pump-probe transmission drop and the experimental magnetic field dependence previously published.[38]

The magnetic effect can be qualitatively explained via the large (T+T) nature of 2^1A_g. In states composed of pairs of triplets, specific triplet pair states under zero-field interaction will have singlet character. The amount of singlet character affects the singlet fission rate, and thus the number of singlets that undergo fluorescence. The application of a magnetic field affects the number and degree of triplet pair states with singlet character, resulting in changes in the rates of fission and fusion. This leads to modulations in the amount of generated fluorescence on the timescale of the fission and fusion events.[5,55,56] The large (T+T) character of 2^1A_g means that a magnetic field effect on the time-resolved fluorescence signal is plausible, but further analysis is outside the scope of this article.

The consistency of a long-lived 2^1A_g with the strong nanosecond pump-probe transmission drop is now examined.[38] This requires evidence that long-lived 2^1A_g states can produce the same magnitude of absorption as required by the reported 176% singlet fission yield.[4] Of the two cases, singlet fission will produce more excitons due to the creation of two independent triplets per photon absorption. Therefore, a long-lived 2^1A_g will require a stronger molar absorptivity to produce the same absorbance. By approximating the transition of singlet fission generated excitons by a $T_1 \rightarrow T_2$ transition, and using experimental parameters combined with Beer's Law, a long-lived 2^1A_g would need a molar absorptivity of 16,600 $M^{-1}cm^{-1}$ compared to the 9,500 $M^{-1}cm^{-1}$ for a $T_1 \rightarrow T_2$ transition.[4] Thus, the $2^1A_g \rightarrow n^1B_u$ transition requires a molar absorptivity 1.75 times, or equivalently a transition dipole moment (TDM) 1.32 times that of the $T_1 \rightarrow T_2$ transition to attain the same magnitude of absorption (SI of ref. 38). Using RAS-SF TDMs (Table 2.2), we see that 2^1B_u and 3^1B_u have TDM ratios of 1.17 and 0.51 respectively. Furthermore, these two states are energetically near one another, being separated by only 0.2 eV (Figure 2.2). Assuming both states are nearly resonant with the probe pulse, we calculate a $2^1A_g \rightarrow n^1B_u$ to $T_1 \rightarrow T_2$ ratio of 1.27 (SI of ref. 38). This ratio is large enough for a $2^1A_g \rightarrow n^1B_u$ transition to account for the observed transmission drop, indicating that the transmission drop is consistent with absorption in the 2^1A_g state.

Table 2.2 Transition moment dipoles, computed at RAS-SF(10,8) with a quintet reference, and their ratios with the $T_1 \rightarrow T_2$ transition.

	$2^1A_g \rightarrow 1^1B_u$	$2^1A_g \rightarrow 2^1B_u$	$2^1A_g \rightarrow 3^1B_u$	$T_1 \rightarrow T_2$
TDM (debye)	2.06	2.96	1.29	2.54
Ratio with $T_1 \rightarrow T_2$ transition	0.81	1.17	0.51	1.00
Transition Energy (eV)	0.37	1.24	1.44	1.74

2.4.3.1 Internal Conversion from $2^1A_g \rightarrow 1^1A_g$

Having demonstrated the consistency of a long-lived 2^1A_g with experimental magnetic field effects and pump-probe transmission results, we examine the longevity of a 2^1A_g exciton. An excited singlet's population can decay through fluorescence, intersystem crossing, or internal conversion. Fluorescence from 2^1A_g to 1^1A_g is symmetry forbidden, (though activated delayed fluorescence from 2^1A_g to 1^1B_u is possible) and we have shown that ISC is negligible due to the calculated low spin-orbit coupling. This leaves internal conversion as the only decay pathway yet uncharacterized as inactive. This is difficult to unequivocally establish without sampling all possible degrees of freedom accessible over microsecond timescales. The rigid structure of QOT2, however, can be shown to hinder internal conversion pathways for 2^1A_g.

To this point, there are currently no studies on the internal conversion dynamics of QOT2. Given QOT2's polyene-like structure, studies on 1,3,5,7-trans-octatetraene can provide some insight into possible QOT2 internal conversion pathways. In octatetraene and QOT2, absorption results in a 1^1B_u exciton that rapidly transitions to a 2^1A_g surface.[57–59] Both systems also display longer excited state lifetimes than expected. Early studies on octatetraene, which were unable to detect the weak 2^1A_g fluorescence, used absorption spectra to calculate a radiative lifetime of 1-2 ns.[60,61] Later studies measured lifetimes of 225 ns at 10 K due to the symmetry-forbidden nature of the 2^1A_g to 1^1A_g transition.[62] Islampour et al. used time- and frequency-domain approaches beyond the Condon approximation to simulate that internal conversion from 2^1A_g to 1^1A_g would occur on a 2 μs time scale for an octatetraene molecule in a vacuum, although inclusion of bath interactions would significantly reduce this value.[63] Other experimental and computational studies suggested a non-adiabatic *trans*→*cis* isomerization as the main non-radiative decay channel through which octatetraene returns to the ground state.[64,65] The isomerization is initiated by a rotation of the octatetraene backbone, a general method of reaching conical intersections in linear polyenes.[66] These studies indicate that rotations along the QOT2 backbone may induce internal conversion.

Due to the rigidity of the thiophene rings, rotations along the QOT2 backbone will most likely occur along either the central C=C or the terminal C=C's at the dicyano termini. The central C=C cis-trans isomerization coordinate is considered first. A RAS-SF(4,4) scan (Figure 2.7) shows that this motion will be resisted, costing a non-negligible 1.20 eV to rotate 90 degrees. Furthermore, neither the 1^1A_g nor the 1^5A_g surfaces come significantly close to contacting the 2^1A_g surface, maintaining a separation of at least 1.25 eV and 2.40 eV respectively. Thus, *cis*→*trans* isomerization along the central C=C does not seem to be a viable pathway for either internal conversion or electronically decoupled $^1(TT)$ formation in QOT2.

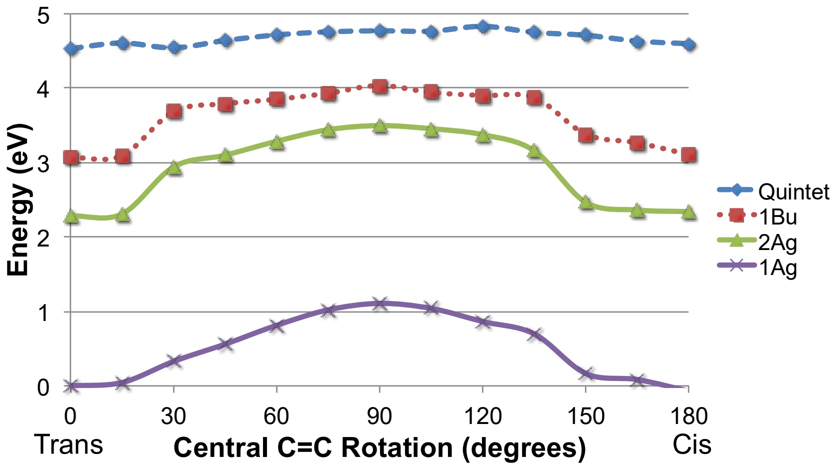


Figure 2.7 RAS-SF(4,4) scan of rotation about the central C=C bond. Salient features include the continuously large 2^1A_g-1^1A_g gap and the 0.75 eV activation barrier to *trans*→*cis* isomerization

The second class of bonds potentially susceptible to rotations are the C=C bonds connecting the dicyano groups to thiophene rings. Since there are two of these bonds, rotations may occur singly or simultaneously. Simultaneous rotation was investigated earlier in pursuit of an independent triplet geometry (Figure 2.6b). Similar to rotation about the central C=C, it was found that this motion results in a significant energy gap of 1.98 eV between 1^1A_g and 2^1A_g, indicating a non-operative internal conversion pathway. Rotation of a single dicyano group (Figure 2.8) also results in a sizeable energy gap of at least 1.6 eV at all points. Thus, an accessible conical intersection along twists of cyano groups appears unlikely.

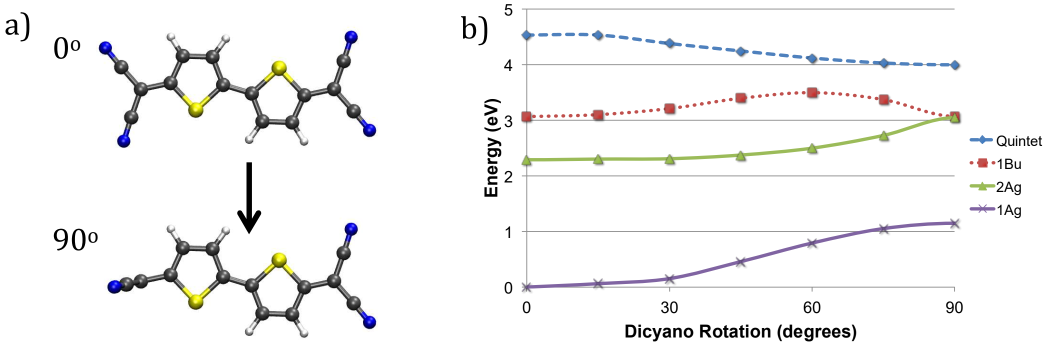


Figure 2.8 a) Geometries of the RAS-SF scanning coordinate. One cyano group is twisted while all else remains frozen. b) RAS-SF(4,4) scan of the rotation of one cyano group starting from the 2^1A_g minimum. Salient features include the consistently large 1^1A_g-2^1A_g gap and the degeneracy between 1^1B_u and 2^1A_g at 90 degrees rotation

The potential energy scans of Figures 2.7 and 2.8 suggest that out-of-plane rotational pathways are inoperative in QOT2 as internal conversion pathways. They also indicate that efficient QOT2 internal conversion, if it occurs, will need to undergo more complicated movements to do so. In sum, the 2^1A_g state is likely long-lived, which is consistent with three central pieces of experimental data (microsecond transient absorption spectra, magnetic dependence, and two-color transmission).[38] We now seek to ascertain whether this remains true for delayed fluorescence measurements.

2.4.3.2 Delayed Fluorescence and Comparison of Singlet Fission Mechanisms

The detection of delayed fluorescence (DF) up to 30 ns at 580nm (2.14 eV), separate from the prompt fluorescence at 470nm (2.64 eV),[38] can be qualitatively explained by a long-lived 2^1A_g. As discussed earlier, passage through the conical intersection results in vibrationally excited 2^1A_g excitons. At early times following excitation, this extra energy will assist a return to the conical intersection and repopulation of the 1^1B_u surface from which delayed fluorescence can occur (Figure 2.9). The difference in energy between prompt and delayed fluorescence results from the differing vibrational populations in the bright 1^1B_u state. Prompt fluorescence occurs from higher 1^1B_u levels due to absorption of pump photons (420 nm, 2.95 eV) with energy in excess of the 1^1A_g-1^1B_u gap (2.69 eV, Figure 2.2). The lower DF energy is a result of the uphill cost of returning to 1^1B_u from 2^1A_g (0.37 eV, Figure 2.2), which results in repopulation of only the lowest 1^1B_u vibrational levels. The decay of the signal is due to dissipation of energy from 2^1A_g vibrational modes over time, rendering them unable to reach the conical intersection. The high DF energy (2.14 eV) means that phosphorescence from T_1 (0.94 eV) cannot account for this signal, providing further evidence against ISC.

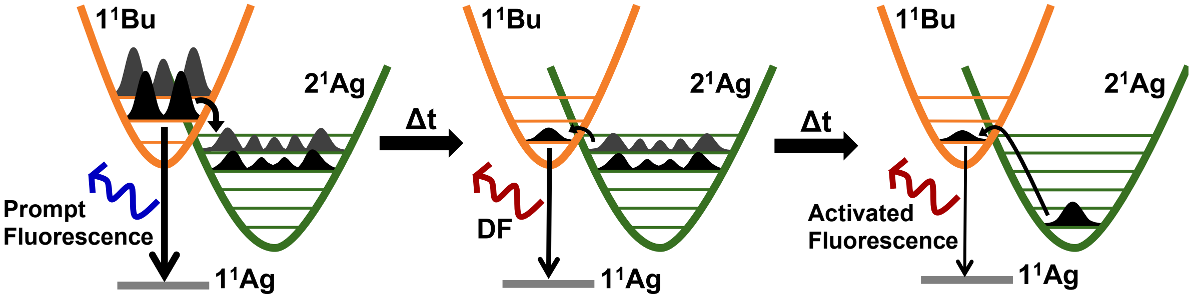


Figure 2.9 Rationalization of experimental DF from the viewpoint of a long-lived 2^1A_g. The leftmost picture depicts the situation after absorption of the pump pulse. The higher vibrational levels of 1^1B_u are populated, and prompt fluorescence occurs from these levels. Population of excited vibrational levels of 2^1A_g occurs via a conical intersection as well. Subsequently, as depicted by the central picture, repopulation of the lower 1^1B_u vibrational levels occurs, leading to DF of a lower energy than the prompt fluorescence. Over time, the 2^1A_g excitons lose vibrational energy via bath interactions, depicted by the rightmost picture. Fluorescence occurring at this stage will be activated, and is expected to be minimal.

The DF signal, while suggestive of the 2^1A_g exciton, is not a direct observation of such a state. Overall, the available experimental data does not continuously track the existence of a 2^1A_g exciton from its formation all the way to the microsecond time range. Specifically, femtosecond pump-probe experiments[4,15] observe the ultrafast formation of a consistent dark state from 1 to 30 ps, and microsecond transient absorption[38] observes the decay of a species from 1 to 57 µs. To link the two timescales, DF measurements[38] are available reaching up to 30 ns. Lacking explicit signals from 30 to 1,000 ns opens the possibility that 2^1A_g evolves into another state before being detected in the microseconds. However, the above investigations of transformations such as ISC and internal conversion give no indication that these changes are favorable.

Finally, we discuss the possibility of alternative singlet fission pathways. Musser et al. have found indications that polyene-like systems can bypass 2^1A_g when undergoing singlet fission.[67,68] For instance, in a poly(thineylene-vinylene), activated intramolecular singlet fission was suggested to occur via internal conversion from a hot 1^1B_u to a n^1A_g state lying above 2^1A_g.[67] This mechanism is unlikely in QOT2 given the two-color transmission experiment's pump energy (445 nm, 2.78 eV) and calculations that place the 3^1A_g state 1.3 eV above 1^1B_u's 2.69 eV. Another alternative singlet fission path was noted in carotenoid aggregates, where it was proposed that the shape of the aggregate PES causes the exciton to access a conical intersection other than the 1^1B_u-2^1A_g conical intersection.[68]

This mechanism is also unlikely in QOT2 due to a lack of detected concentration dependence for the steady state spectra, indicating negligible aggregate formation.[4]

The inapplicability of these alternative singlet fission mechanisms indicates that for triplet decoupling to occur in QOT2's ^{1}(TT) state, a geometrical deformation is required. However, the motion determined above—simultaneous twist of the two dicyano groups—is energetically unfavorable. In fact, investigations into a number of rotational motions showed that, generally speaking, triplet decoupling in ^{1}(TT) would not occur through rotations of the QOT2 backbone. The singlet fission process in QOT2 is therefore likely to end without full triplet decoupling, but with significant biexcitonic character arising from the large (T+T) component of ^{1}(TT) shown in Figure 2.5.

2.5 Conclusion

The presented computational studies provide detailed information on the character and time evolution of QOT2 excited states. Quantum dynamics simulations suggest that ultrafast decay, mediated by a few key vibrations, occurs from the bright 1^1B_u state to the dark 2^1A_g state through a conical intersection located near the Franck-Condon region. This result agrees with experimental observation of ultrafast (<1 ps) decay of the bright 1^1B_u, and shows that 2^1A_g is rapidly populated following photo excitation.

Two main avenues of 2^1A_g evolution were considered in the course of this work, with the first being intramolecular singlet fission. This investigation resulted in the first detailed characterization of the QOT2 2^1A_g state, which showed that planar 2^1A_g is a strongly coupled ^{1}(TT) state where the two triplets incompletely localize to each half of the molecule. Completion of singlet fission by spatially separating and electronically decoupling the two triplets in 2^1A_g is unlikely due to a large cost of 1.76 eV for reaching a twisted geometry capable of hosting two independent triplets. A second possibility, that the observed long-lived excited state is the multiexcitonic 2^1A_g state, is supported by TDM calculations that show sufficient oscillator strength around the necessary absorption region to account for the observed transitions. The longevity of the 2^1A_g exciton is further supported by the computed low spin-orbit coupling to the triplet manifold and low activity of various internal conversion pathways. While the presented evidence does not allow for a conclusive assignment of the μs transient absorption signal to a long-lived 2^1A_g, explicitly tracking the 2^1A_g exciton from its formation through the microsecond time scale would resolve this question. These investigations into intramolecular singlet fission and a long-lived 2^1A_g are summarized in Table 2.3.

Table 2.3 Summary of investigations of intramolecular singlet fission and a long-lived 2^1A_g.

Intramolecular singlet fission	Long-Lived 2^1A_g
+ 2^1A_g rapidly generated by conical intersection (Section 2.4.1) + 2^1A_g characterized as a ^{1}(TT) state (Section 2.4.2.2) + Geometry suitable for separated triplets found by constrained HF (Section 2.4.2.2) – Large barrier (1.76 eV) for reaching separated triplet geometry (Section 2.4.2.2) – Rotational scans and decoupled triplet optimization find no other geometries for triplet separation (Section 2.4.3.1)	+ 2^1A_g rapidly generated by conical intersection (Section 2.4.1) + Low spin-orbit coupling indicates negligible ISC (Section 2.4.2.1) + Magnetic dependence of emission accounted for by 2^1A_g ^{1}(TT) character (Section 2.4.3) + Transient absorption signals accounted for by TDM calculations on 2^1A_g (Section 2.4.3) + Sampling of internal conversion pathways found to be inactive due to rigid QOT2 structure (Section 2.4.3.1) – No experimental confirmation this state exists in μs timescale (Section 2.4.3.2)

The large barrier for triplet separation in QOT2 indicates that ^{1}(TT) decoupling may be a significant challenge in intramolecular singlet fission systems composed of fully connected, conjugated fragments. It remains to be seen if strongly coupled ^{1}(TT) states can be efficiently extracted as multiple charge carriers, negating the need for full triplet separation. In contrast, systems with distinct electronic subunits, such as covalently linked acene dimers,[50,51] likely will not experience this same challenge. Whether intramolecular singlet fission chromophores require structural motifs that allow facile pathways for efficient triplet separation therefore will depend on the efficacy of multi-charge extraction.

In sum, this contribution provides the first physical description of the dark, highly coupled, double triplet ^{1}(TT) state in QOT2, revealing that the coupling results from spatially overlapping localized orbitals. Intriguingly, this state has an excited state absorption signature similar to that of two triplets, despite the strong coupling and spatial inseparability of the excitons. Ultrafast formation of the ^{1}(TT) state mediated by key vibrational modes suggests that multiexcitonic states can be rapidly and efficiently generated through conical intersections in monomeric systems.

While this Chapter shows that RAS-SF methods can provide key insights on the relative energies of multi-excitonic states, it sidesteps one of the weaknesses of current RAS-SF implementations. RAS(h,p)-SF has a poor treatment of dynamic correlation due to the small amount of excitations that are included outside of the active space, resulting in poor energetics for excitonic states. This problem is avoided here by utilizing XMS-CASPT2 calculations for these states. Chapter 3 endeavors to increase the general applicability of RAS-SF methods by increasing the number of excitations outside of the active space.

2.6 References

1. Shockley, W. & Queisser, H. J. Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *J. Appl. Phys.* **32,** 510–519 (1961).
2. Hanna, M. C. & Nozik, A. J. Solar conversion efficiency of photovoltaic and photoelectrolysis cells with carrier multiplication absorbers. *J. Appl. Phys.* **100,** 74510 (2006).
3. Tayebjee, M. J. Y., Gray-Weale, A. A. & Schmidt, T. W. Thermodynamic Limit of Exciton Fission Solar Cell Efficiency. *J. Phys. Chem. Lett.* **3,** 2749–2754 (2012).
4. Varnavski, O. *et al.* High Yield Ultrafast Intramolecular Singlet Exciton Fission in a Quinoidal Bithiophene. *J. Phys. Chem. Lett.* **6**, 1375–1384 (2015).
5. Smith, M. B. & Michl, J. Singlet fission. *Chem. Rev.* **110,** 6891–936 (2010).
6. Smith, M. B. & Michl, J. Recent advances in singlet fission. *Annu. Rev. Phys. Chem.* **64,** 361–86 (2013).
7. Ponce Ortiz, R. *et al.* On the biradicaloid nature of long quinoidal oligothiophenes: Experimental evidence guided by theoretical studies. *Angew. Chemie - Int. Ed.* **46,** 9057–9061 (2007).
8. Casado, J., Ponce Ortiz, R. & López Navarrete, J. T. Quinoidal oligothiophenes: new properties behind an unconventional electronic structure. *Chem. Soc. Rev.* **41,** 5672–5686 (2012).
9. Ponce Ortiz, R. *et al.* Quinoidal Oligothiophenes: Towards Biradical Ground-State Species. *Chem. - A Eur. J.* **16,** 470–484 (2010).
10. Casado, J. & López Navarrete, J. T. The longest quinoidal oligothiophene: A Raman story. *Chem. Rec.* **11,** 45–53 (2011).
11. Fazzi, D., Canesi, E. V., Negri, F., Bertarelli, C. & Castiglioni, C. Biradicaloid character of thiophene-based heterophenoquinones: The role of electron-phonon coupling. *ChemPhysChem* **11,** 3685–3695 (2010).
12. Paci, I. *et al.* Singlet fission for dye-sensitized solar cells: Can a suitable sensitizer be found? *J. Am. Chem. Soc.* **128,** 16546–16553 (2006).

13. Minami, T. & Nakano, M. Diradical Character View of Singlet Fission. *J. Phys. Chem. Lett.* **3,** 145–150 (2012).
14. Minami, T., Ito, S. & Nakano, M. Fundamental of Diradical-Character-Based Molecular Design for Singlet Fission. *J. Phys. Chem. Lett.* **4,** 2133–2137 (2013).
15. Wang, Z. & Kobayashi, T. Electronic relaxation and vibrational dynamics in a thiophene oligomer studied under the same experimental condition with a sub-5 fs laser. *New J. Phys.* **10,** 123021–123034 (2008).
16. Di Motta, S., Negri, F., Fazzi, D., Castiglioni, C. & Canesi, E. V. Biradicaloid and polyenic character of quinoidal oligothiophenes revealed by the presence of a low-lying double-exciton state. *J. Phys. Chem. Lett.* **1,** 3334–3339 (2010).
17. Hosteny, R. P. Ab initio study of the π-electron states of trans-butadiene. *J. Chem. Phys.* **62,** 4764–4779 (1975).
18. Schulten, K. & Karplus, M. On the origin of a low-lying forbidden transition in polyenes and related molecules. *Chem. Phys. Lett.* **14,** 305–309 (1972).
19. Greyson, E. C., Vura-Weis, J., Michl, J. & Ratner, M. A. Maximizing singlet fission in organic dimers: theoretical investigation of triplet yield in the regime of localized excitation and fast coherent electron transfer. *J. Phys. Chem. B* **114,** 14168–77 (2010).
20. Feng, X., Luzanov, A. V. & Krylov, A. I. Fission of Entangled Spins: An Electronic Structure Perspective. *J. Phys. Chem. Lett.* **4,** 3845–3852 (2013).
21. Kolomeisky, A. B., Feng, X. & Krylov, A. I. A Simple Kinetic Model for Singlet Fission: A Role of Electronic and Entropic Contributions to Macroscopic Rates. *J. Phys. Chem. C* **118,** 5188–5195 (2014).
22. Matsika, S., Feng, X., Luzanov, A. V & Krylov, A. I. What We Can Learn from the Norms of One-Particle Density Matrices, and What We Can't: Some Results for Interstate Properties in Model Singlet Fission Systems. *J. Phys. Chem. A* **118,** 11943–11955 (2014).
23. Luzanov, A. V., Casanova, D., Feng, X. & Krylov, A. I. Quantifying charge resonance and multiexciton character in coupled chromophores by charge and spin cumulant analysis. *J. Chem. Phys.* **142,** 224104 (2015).
24. Zimmerman, P. M., Bell, F., Goldey, M., Bell, A. T. & Head-Gordon, M. Restricted active space spin-flip configuration interaction: Theory and examples for multiple spin flips with odd numbers of electrons. *J. Chem. Phys.* **137,** 164110 (2012).
25. Zimmerman, P. M., Zhang, Z. & Musgrave, C. B. Singlet fission in pentacene through multi-exciton quantum states. *Nat. Chem.* **2,** 648–52 (2010).
26. Zimmerman, P. M., Bell, F., Casanova, D. & Head-Gordon, M. Mechanism for Singlet Fission in Pentacene and Tetracene: From Single Exciton to Two Triplets. *J. Am. Chem. Soc.* **133,** 19944–19952 (2011).
27. Zimmerman, P. M., Musgrave, C. B. & Head-Gordon, M. A Correlated Electron View of Singlet Fission. *Acc. Chem. Res.* **46,** 1339–1347 (2013).
28. Musser, A. J. *et al.* Evidence for conical intersection dynamics mediating ultrafast singlet exciton fission. *Nat. Phys.* **11,** 352–357 (2015).
29. Izumi, T., Kobashi, S., Takimiya, K., Aso, Y. & Otsubo, T. Synthesis and spectroscopic properties of a series of β-blocked long oligothiophenes up to the 96-mer: re-evaluation of effective conjugation length. *J. Am. Chem. Soc.* **125,** 5286–5287 (2003).
30. Takahashi, T., Matsuoka, K. I., Takimiya, K., Otsubo, T. & Aso, Y. Extensive quinoidal oligothiophenes with dicyanomethylene groups at terminal positions as highly

amphoteric redox molecules. *J. Am. Chem. Soc.* **127,** 8928–8929 (2005).

31. Shiozaki, T., Győrffy, W., Celani, P. & Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **135,** 81106 (2011).
32. Werner, H.-J., Knowles, P. J., Knizia, G., Manby, F. R. & Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2,** 242–253 (2012).
33. Meyer, H.-D., Manthe, U. & Cederbaum, L. S. The multi-configurational time-dependent Hartree approach. *Chem. Phys. Lett.* **165,** 73–78 (1990).
34. Beck, M. The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. *Phys. Rep.* **324,** 1–105 (2000).
35. Worth, G., Meyer, H.-D., Jäckle, A. & Beck, M. The Heidelberg MCTDH Package. Available at: www.pci.uni-heidelberg.de/tc/usr/mctdh/doc.85/index.html. (Accessed: 23rd September 2015)
36. Mikhailov, I. A., Tafur, S. & Masunov, A. E. Double excitations and state-to-state transition dipoles in π-π* excited singlet states of linear polyenes: Time-dependent density-functional theory versus multiconfigurational methods. *Phys. Rev. A* **77,** 12510 (2008).
37. Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **113,** 184–215 (2015).
38. Chien, A. D. *et al.* Structure and Dynamics of the ^{1}(TT) State in a Quinoidal Bithiophene: Characterizing a Promising Intramolecular Singlet Fission Candidate. *J. Phys. Chem. C* **119,** 28258–28268 (2015).
39. Rossi, R., Ciofalo, M., Carpita, A. & Ponterini, G. Singlet—triplet intersystem crossing in 2,2′:5′,2″-terthiophene and some of its derivatives. *J. Photochem. Photobiol. A Chem.* **70,** 59–67 (1993).
40. Paa, W., Yang, J.-P., Helbig, M., Hein, J. & Rentsch, S. Femtosecond time-resolved measurements of terthiophene: fast singlet–triplet intersystem crossing. *Chem. Phys. Lett.* **292,** 607–614 (1998).
41. Becker, R. S., Demelo, J. S., Macanita, A. L. & Elisei, F. Comprehensive evaluation of the absorption, photophysical, energy transfer, structural, and theoretical properties of alpha-oligothiophenes with one to seven rings. *J. Phys. Chem.* **100,** 18683–18695 (1996).
42. Beljonne, D., Shuai, Z., Pourtois, G. & Bredas, J. L. Spin - Orbit Coupling and Intersystem Crossing in Conjugated Polymers: A Configuration Interaction Description. *J. Phys. Chem. A* **105,** 3899–3907 (2001).
43. Siegert, S., Vogeler, F., Marian, C. M. & Weinkauf, R. Throwing light on dark states of α-oligothiophenes of chain lengths 2 to 6: radical anion photoelectron spectroscopy and excited-state theory. *Phys. Chem. Chem. Phys.* **13,** 10350–10363 (2011).
44. Sun, S., Zhang, S., Liu, K., Wang, Y. & Zhang, B. The geometry relaxation and intersystem crossing of quaterthiophene studied by femtosecond spectroscopy. *Photochem. Photobiol. Sci.* **14,** 853–858 (2015).
45. Nijegorodov, N., Zvolinsky, V. & Luhanga, P. V. C. Photonics and photochemical stability of aromatic molecules, family related in π-structure but different in planarity, rigidity and molecule symmetry. *J. Photochem. Photobiol. A Chem.* **196,** 219–226 (2008).

46. El-Sayed, M. A. Spin—Orbit Coupling and the Radiationless Processes in Nitrogen Heterocyclics. *J. Chem. Phys.* **38,** 2834–2838 (1963).
47. McClure, D. S. Selection Rules for Singlet-Triplet Perturbations in Polyatomic Molecules. *J. Chem. Phys.* **17,** 665–666 (1949).
48. Weissman, S. I. Vector Model for Spin-Orbit Interaction in Polyatomic Molecules. *J. Chem. Phys.* **18,** 232–233 (1950).
49. Marcus, R. A. Electron Transfer Reactions in Chemistry: Theory and Experiment (Nobel Lecture). *Angew. Chemie Int. Ed. English* **32,** 1111–1121 (1993).
50. Müller, A. M., Avlasevich, Y. S., Schoeller, W. W., Müllen, K. & Bardeen, C. J. Exciton fission and fusion in bis(tetracene) molecules with different covalent linker structures. *J. Am. Chem. Soc.* **129,** 14240–14250 (2007).
51. Zirzlmeier, J. *et al.* Singlet fission in pentacene dimers. *Proc. Natl. Acad. Sci.* **112,** 5325–5330 (2015).
52. Busby, E. *et al.* A design strategy for intramolecular singlet fission mediated by charge-transfer states in donor–acceptor organic materials. *Nat. Mater.* **14,** 426–433 (2015).
53. Gradinaru, C. C. *et al.* An unusual pathway of excitation energy deactivation in carotenoids: Singlet-to-triplet conversion on an ultrafast timescale in a photosynthetic antenna. *Proc. Natl. Acad. Sci.* **98,** 2364–2369 (2001).
54. Papagiannakis, E. *et al.* Light Harvesting by Carotenoids Incorporated into the B850 Light-Harvesting Complex from Rhodobacter s phaeroides R-26.1: Excited-State Relaxation, Ultrafast Triplet Formation, and Energy Transfer to Bacteriochlorophyll. *J. Phys. Chem. B* **107,** 5642–5649 (2003).
55. Burdett, J. J. & Bardeen, C. J. Quantum beats in crystalline tetracene delayed fluorescence due to triplet pair coherences produced by direct singlet fission. *J. Am. Chem. Soc.* **134,** 8597–607 (2012).
56. Burdett, J. J., Piland, G. B. & Bardeen, C. J. Magnetic field effects and the role of spin states in singlet fission. *Chem. Phys. Lett.* **585,** 1–10 (2013).
57. Yee, W. A., O'Neil, R. H., Lewis, J. W., Zhang, J. Z. & Kliger, D. S. Femtosecond Studies of Electronic Relaxation, Vibrational Relaxation, and Rotational Diffusion in all -trans-1,8-Diphenyl-1,3,5,7-octatetraene. *J. Phys. Chem. A* **103,** 2388–2393 (1999).
58. Ohta, K., Naitoh, Y., Tominaga, K. & Yoshihara, K. Excited-State Dynamics of all-trans-1,3,5,7-Octatetraene in Solution. Direct Observation of Internal Conversion from the S2 to S1 State and Relaxation Processes in the S1 State. *J. Phys. Chem. A* **105,** 3973–3980 (2001).
59. Islampour, R., Khavaninzadeh, A. & Miralinaghi, M. Theoretical analysis of vibronic structure in absorption and fluorescence spectra of polyatomic molecules beyond the Condon approximation: Application to $1^1A_g \leftrightarrow 2^1A_g$ and $1^1A_g \leftrightarrow 1^1B_u$ electronic transitions in all-trans-1,3,5,7-octatetraene. *J. Mol. Spectrosc.* **286–287,** 30–45 (2013).
60. Gavin, R. M., Weisman, C., McVey, J. K. & Rice, S. A. Spectroscopic properties of polyenes. III. 1,3,5,7-Octatetraene. *J. Chem. Phys.* **68,** 522–529 (1978).
61. Kohler, B. E. Octatetraene photoisomerization. *Chem. Rev.* **93,** 41–54 (1993).
62. Ackerman, J. R. Radiationless decay of 1,3,5,7-octatetraene. *J. Chem. Phys.* **77,** 3967–3973 (1982).
63. Islampour, R., Miralinaghi, M. & Khavaninzadeh, A. Excited-State Dynamics of trans ,

trans -1,3,5,7-Octatetraene: Estimation of Decay Rate Constants of 1^1B_u -> 2^1A_g and 2^1A_g -> 1^1A_g Internal Conversions. *J. Phys. Chem. A* **115,** 8860–8869 (2011).

64. Petek, H. *et al.* The 2^1A_g state of trans,trans-1,3,5,7-octatetraene in free jet expansions. *J. Chem. Phys.* **98,** 3777–3794 (1993).
65. Garavelli, M. *et al.* The Structure of the Nonadiabatic Photochemical Trans → Cis Isomerization Channel in All-Trans Octatetraene. *J. Am. Chem. Soc.* **118,** 11656–11657 (1996).
66. Celani, P. *et al.* Molecular 'Trigger' for the Radiationless Deactivation of Photoexcited Conjugated Hydrocarbons. *J. Am. Chem. Soc.* **117,** 11584–11585 (1995).
67. Musser, A. J. *et al.* Activated singlet exciton fission in a semiconducting polymer. *J. Am. Chem. Soc.* **135,** 12747–12754 (2013).
68. Musser, A. J. *et al.* The Nature of Singlet Exciton Fission in Carotenoid Aggregates. *J. Am. Chem. Soc.* **137,** 5130–5139 (2015).

Chapter 3: Recovering Dynamic Correlation in Spin Flip Configuration Interaction through a Difference Dedicated Approach

This Chapter largely based upon published work:
Reproduced with permission from Chien, A. D. & Zimmerman, P. M. *J. Chem. Phys.* **146,** 14103 (2017).

3.1 Abstract

This Chapter introduces the restricted-active-space *n*-spin flip configuration interaction models, RAS(S)-SF and RAS(S,2h,2p)-SF, which extends the RAS(h,p)-SF method used in the prior Chapter. This extension aims to provide highly correlated, yet low cost, approaches for treating polyradical systems by adding electronic degrees of freedom beyond those of previous spin flip approaches in order to achieve accurate ground and excited state energetics. The effects of additional dynamic correlation were investigated by comparing these two techniques to the prior RAS(h,p)-SF method on a variety of test systems, including multiple electronic states of methylene, tetramethyleneethane, three binuclear transition metal complexes, and a tetracene dimer. RAS(S,2h,2p)-SF significantly improves state descriptions in all cases and provides high accuracy even when using a minimal number of spin flips. Furthermore, this correlated level of theory is shown to be extensible to the large systems involved in singlet fission, where the multi-excitonic states in tetracene dimers are difficult to simulate with standard methods and therefore are still a matter of debate. Using a triple-zeta basis, the double triplet state, ^{1}(TT), is predicted to be unbound. This result contradicts lower levels of theory and provides important insight into tetracene's ability to undergo efficient singlet fission.

3.2 Introduction

Accurate electronic structure simulations on systems that are well described by a single electron configuration are routinely possible because the complexities associated with treating strong electron correlation are avoided. Situations that contain strong correlations among multiple electronic configurations, including bond breakage, radicals, and excited states, usually require multi-reference techniques such as the widely used complete active space self-consistent field (CASSCF) method.[1,2] CASSCF, however, is not always suited to routine use, as it usually demands prior knowledge of important electron configurations and scales exponentially with increasing active space size. Recently, new approaches have appeared that reduce the cost of CAS calculations, enabling the computation of larger active spaces. Notable examples include the density matrix renormalization group (DMRG)[3,4] and various stochastic approaches,[5–7] which greatly reduce computational cost prefactors for CAS and full configuration interaction (FCI). Select CI approaches have also been recently re-introduced,[8–12] which also greatly reduce the overall cost of electronic structure simulations of strongly correlated systems.

Another alternative is the spin flip (SF) methodology, first introduced by Krylov, which is able to capture strong correlations with a single-reference ansatz based on a high-spin reference.[13–15] By performing spin-flipping excitations on this reference, low-spin electronic configurations are generated to construct CI wave functions. The high-spin restricted open-shell Hartree Fock (ROHF) solution is a natural reference for common multi-reference situations such as bond-breaking or multi-radicals since target states are reasonably well described by a small set of unit-occupied orbitals. For example, the triplet ROHF solution has two unit-occupied frontier orbitals, and therefore acts as a good approximation to a broken single bond or a singlet biradical state. This concept extends to higher order bond-breakages and multi-radicals where higher-spin references such as quintets, septets, and so on are utilized.

The SF methodology therefore can describe strong correlation while avoiding issues that trouble multi-reference methods. For instance, the use of a single ROHF reference eliminates the expense of multiconfigurational orbital optimization, as well as complications arising from the nonlinear nature of this optimization problem.[16]

Furthermore, multi-reference results are typically heavily dependent on the chosen active space, requiring one to approach these calculations with either prior knowledge of the states of interest or careful examination of their active space dependencies. SF methods bypass these issues since the active space is chosen as the unit-occupied orbitals of the high-spin reference.

SF methods also greatly benefit from the balanced nature of the high-spin reference, in that ground and excited state descriptions are treated on a similar footing. To understand this effect, consider that while a RHF reference is a reasonable description of a closed-shell ground state, it is poorly suited for describing excited states. Thus, at any truncated level of CI, an imbalance exists between descriptions of the ground and excited states, and excitation energies are typically overestimated (e.g. CISD[17–19]). ROHF references, on the other hand, offset this imbalance by providing an improved description of low-lying excited states at the expense of the ground state. This results in a balanced handling of ground and excited states at a given CI excitation level, resulting in more accurate relative energies.

These advantages have led to the continual development of SF configuration interaction methods beyond the original technique, spin flip configuration interaction singles (SF-CIS).[14] Within the CI framework, SF-CIS was extended to spin-complete spin flip CIS (SC-SF-CIS), which eliminated spin impurities by adding in spin-complementing configurations.[20] Spin flip extended CIS (SF-XCIS) then added in excitations whose absences resulted in an imbalanced treatment of HOMO and LUMO orbitals.[21] SF-XCIS allows the configuration basis to be viewed as a restricted active space (RAS)[1,22,23] method involving all possible configurations in the active space as well as single excitations into and out of the active space. The most recent SF-CI methods, RAS(h,p)-2SF and RAS(h,p)-nSF, extended the SF-CI family to multiple spin flips,[24–26] enabling the low-cost, balanced treatment of ground and excited states with multi-radical character in large systems.

SF-CI energy gaps for single exciton states, despite the balanced starting point, are semiquantitative due to the low level of included dynamic correlation. Accordingly, dynamically correlated perturbative[27,28] and coupled-cluster[13,29,30,32] SF variants have been developed. Treating dynamic correlation to a higher level may also be accomplished by adding doubles excitations to RAS(h,p)-SF. Doing so has the advantage of guaranteeing spin-purity and being independent of empirical level-shifts, in contrast to coupled-cluster and perturbative methods, respectively. The resulting rise in cost, however, requires that doubles excitations be added with care.

A new SF-CI level of theory that features increased accuracies while preserving modest costs can be obtained by adding only a specific subset of doubles excitations to RAS(h,p)-SF. This idea first gained traction in difference-dedicated configuration interaction methods (DDCI),[32,33] where perturbation theory was used to demonstrate that a certain subset of doubles excitations do not contribute to the energy gaps between states. These doubles are therefore *a priori* removed from the configuration basis. Moreover, the excluded doubles excitations are the largest subset of doubles, allowing DDCI methods to recover dynamic correlation with only a modest increase in computational cost. Adding doubles via DDCI to RAS(h,p)-SF is expected to lead to more quantitative results, especially when dynamic correlation plays a large role in accurate state descriptions.

Herein, the RAS(S,2h,2p)-SF method, which inherits the positive aspects of the RAS(h,p)-SF method while recovering a significant portion of dynamic correlation, is introduced. RAS(S)-SF is also introduced, which includes singles excitations otherwise missing from RAS(h,p)-SF. Benchmarks on a variety of systems at RAS(h,p)-SF, RAS(S)-SF, and RAS(S,2h,2p)-SF levels were performed to determine when dynamic correlation affects energy gaps. As the first examples, methylene (CH_2) and tetramethyleneethane (TME) are investigated as challenging systems with substantial biradical character. Calculations of magnetic coupling in a series of transition metal complexes were also performed to demonstrate applicability to higher levels of radicals. Finally, investigation of a tetracene dimer provides the most sophisticated calculations on a model tetracene singlet fission system to date, demonstrating that dynamic correlation is vital to describe the triplet separation process.

3.3 Theoretical Background of Proposed Spin-Flip Extensions

All of the methods central to this work (DDCI, RAS(h,p)-SF, RAS(S)-SF, and RAS(S,2h,2p)-SF) utilize the RAS framework to define their configuration basis. In the RAS framework, non-frozen orbitals are split into three subsets: RAS2, an active space comprised of both occupied and virtual orbitals; RAS1, the occupied orbitals below RAS2; and RAS3, the virtual orbitals above RAS2. Excitations are classified by the number of holes (empty spaces in RAS1) and particles (electrons in RAS3) generated – i.e. *h-p* excitations create one hole and one particle, 2*h* creates two holes, 2*p* creates two particles, etc. A brief review of the RAS(h,p)-SF and DDCI methods is given before detailing the new RAS(S)-SF and RAS(S,2h,2p)-SF methods.

3.3.1 RAS(h,p)-SF

RAS(h,p)-SF combines a high-spin reference with the configuration basis selection scheme of RAS-CI.[22,23] RAS(h,p)-SF wave functions are expressed by Eq. 3.1, where it can be seen that there are three classes of determinants in the configuration basis. First, any arrangement of the *N* electrons in the *M* orbitals of RAS2 is allowed, i.e. CAS(*N,M*)-CI.

$$\Psi_{RAS(h,p)-SF} = \sum_{\substack{m\in CAS\\(N,M)}} c_m \phi_m + \sum_{h\in RAS1} \sum_{\substack{m\in CAS\\(N+1,M)}} c_m^h \phi_m^h + \sum_{p\in RAS3} \sum_{\substack{m\in CAS\\(N-1,M)}} c_m^p \phi_m^p$$

Equation 3.1 RAS(h,p)-SF wave function

Second, single *h* excitations from RAS1→RAS2 are allowed, in conjunction with any arrangement of (*N*+1) electrons in the *M* orbitals of RAS2. Finally, single *p* excitations from RAS2→RAS3 are allowed, in conjunction with any arrangement of (*N*-1) electrons in the RAS2 orbitals. RAS(h,p)-SF has a number of desirable properties including low-cost, multistate calculations, and spin-pure, variational wave functions. Size-consistency is achieved provided the active space is localized on a single fragment. Additionally, provided the active space stays constant, the configuration basis dimension increases only linearly with system size. The main shortcoming of RAS(h,p)-SF is its inability to quantitatively describe excitation energies of certain classes of excited states due to lack of dynamic correlation.

3.3.2 DDCI

DDCI is another form of RAS-CI that includes CAS-CI within RAS2 space as well as up to doubles excitations into, out of, and around RAS2.[32–37] To maintain low cost, DDCI excludes the *2h-2p* set of excitations from the configuration basis. This choice originates from examination of second-order quasi-degenerate perturbation theory corrections to the active space Hamiltonian, where it is seen that *2h-2p* excitations shift the total energies for all active space states by the same amount and thus do not alter the energy gaps.

DDCI is well-suited for vertical energy gaps, but adiabatic excitation energies are influenced by differences in the discarded doubles excitations between geometries. To account for this, the difference in correlation energy between geometries can be obtained at the level of second-order Moller-Plesset perturbation theory and then applied to the vertical energy gaps.[32,33]

3.3.3 RAS(S)-SF and RAS(S,2h,2p)-SF

Treating dynamic correlation in RAS-SF methods at the DDCI level allows for the following excitations beyond the active space configurations: *h*, *p*, *h-p*, *2h*, *2p, 2h-p*, and *h-2p*. As stated earlier, RAS(h,p)-SF adds *h* and *p* excitations to the configurations obtained by CAS-CI in RAS2. By adding *h-p* excitations to RAS(h,p)-SF, RAS(S)-SF is obtained and designated with an "S" since at this point all *single* excitations into, out of, or around RAS2 have been included. Note that, for spin-completeness, *h-p* excitations include both single excitations from RAS1→RAS3 as well as the formally double (RAS1→RAS2 + RAS2→RAS3) excitations. With RAS(S,2h,2p)-SF, *2h* and *2p* excitations into and out of the active space are included on top of the "S" set of excitations. RAS(S)-SF and RAS(S,2h,2p)-SF wave functions are expressed as in Eq. 3.2 and 3.3, using the same notation as in Eq. 3.1. The configuration basis of all the RAS-SF methods is also graphically depicted in Figure 3.1.

$$\Psi_{RAS(S)-SF} = \Psi_{RAS(h,p)-SF} + \sum_{h\in RAS1} \sum_{p\in RAS3} \sum_{\substack{m\in CAS \\ (N,M)}} c_m^{hp} \phi_m^{hp}$$

Equation 3.2 RAS(S)-SF wave function

$$\Psi_{RAS(S,2h,2p)-SF} = \Psi_{RAS(S)-SF} + \sum_{h_1,h_2\in RAS1} \sum_{\substack{m\in CAS\\(N+2,M)}} c_m^{h_1h_2}\phi_m^{h_1h_2} + \sum_{p_1,p_2\in RAS3} \sum_{\substack{m\in CAS\\(N-2,M)}} c_m^{p_1p_2}\phi_m^{p_1p_2}$$

Equation 3.3 RAS(S,2h,2p)-SF wave function

A method including *2h-p* and *h-2p* excitations was not implemented, but represents the next set of excitations that could improve the accuracy of the RAS-SF family of methods. RAS(S,2h,2p)-SF already provides substantial improvements over RAS(h,p)-SF, providing some numerical evidence that higher excitation levels may not be necessary in many cases.

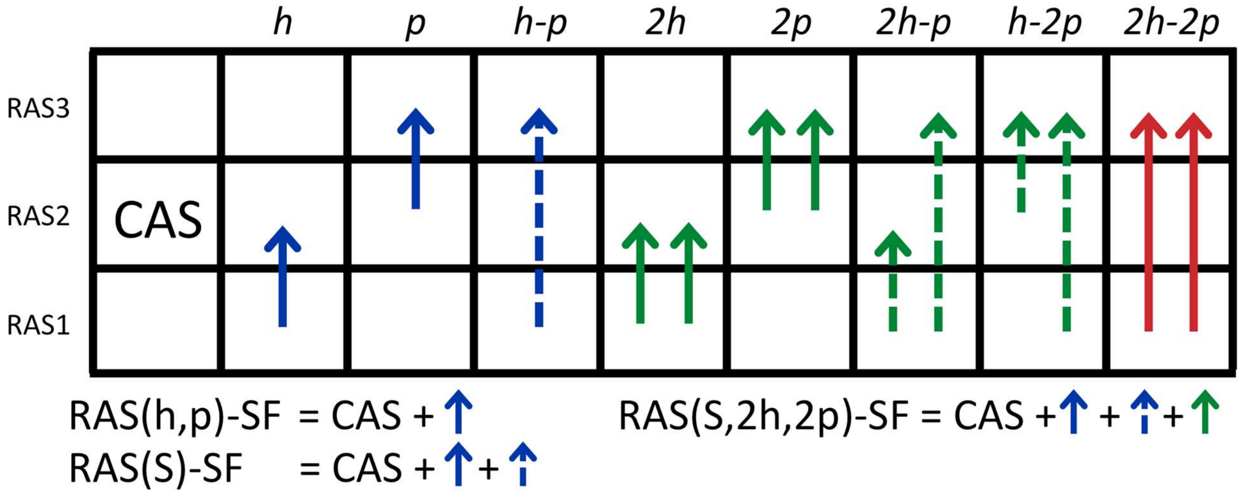


Figure 3.1 Schematic illustrating the RAS-SF methods. The blue singles excitations involve either α or β electrons. Similarly, the green doubles excitations span three combinations of electron spins (two α, two β, or one α and one β). The red doubles are the set discarded by DDCI theory.

3.3.3.1 Size-Extensivity/Intensivity

Size-extensivity is attained when the ground state energy of two non-interacting fragments equals the sum of each individual fragment's energy ($E_{AB} = E_A + E_B$). Size-consistency is a special case of size-extensivity that implies a correct description of chemical reactions such as bond dissociation.[38] Previous analyses of RAS(h,p)-SF showed it to be size-extensive and consistent, provided that the active space was localized on a single fragment.[14,24,25] RAS(S,2h,2p)-SF retains this property as none of the newly added excitations (*h-p*, *2h*, *2p*) are able to relax the ground states of non-interacting fragments.

A similar concept exists for excitation energies, when the addition of a non-interacting fragment B does not affect the excitation energies of the fragment of interest A, termed size-intensivity.[38] Following the analysis of Krylov,[14] it is clear that RAS(S,2h,2p)-SF is size-intensive given an active space localized to fragment A. The newly added single *h-p* excitations produce size-intensive excitation energies at the configuration interaction singles level on the non-interacting fragment B as well.

3.3.3.2 General Properties

Like most CI methods, RAS(S,2h,2p)-SF is invariant to orbital rotations in specific subspaces. Specifically, RAS(S,2h,2p)-SF is invariant to orbital rotations within the RAS1, RAS2, and RAS3 spaces, but variant to interspace rotations. RAS(S,2h,2p)-SF computations solve for multiple, variational, spin-pure states in a single calculation. Spin-purity is not guaranteed in the coupled-cluster family of spin flip methods,[13,30,31] so RAS(S,2h,2p)-SF is the only SF method correlated to the level of doubles that is spin-pure.

3.3.3.3 Implementation

The RAS(S,2h,2p)-SF implementation follows RAS(h,p)-SF,[25] where the configuration basis is composed of Slater determinants, as in the work of Handy.[39,40] Each determinant is uniquely defined by its RAS2 occupations, explicitly stored as α and β occupation strings, and hole/particle indices as necessary. Hamiltonian matrix construction utilizes the fact that determinants more than two excitations away from each other do not contribute to the Hamiltonian matrix, as stated by Slater's rules.[41] This allows the Hamiltonian to be constructed by finding all connections of a certain type - either none (e.g., the diagonal elements), α, β, αα, ββ, or αβ excitations - and computing their contribution to the Hamiltonian matrix. In this way, recognizing that certain RAS(S,2h,2p)-SF determinants cannot be connected by certain types of excitations significantly lowers the computational cost.

Storing and diagonalizing the entire Hamiltonian becomes prohibitively expensive for larger systems. Therefore, determination of the lowest roots of the Hamiltonian uses the Davidson algorithm,[42,43] a subspace iterative diagonalization method. This algorithm requires only the storage of σ vectors, which are contractions of the Hamiltonian with CI vectors, as opposed to the full Hamiltonian. The Davidson algorithm requires a starting guess for the eigenvectors of interest, which is supplied by explicit diagonalization of a guess Hamiltonian that has more frozen RAS1 and RAS3 orbitals, but the same RAS2 space. As the majority of work in Davidson is in the formation of the σ vectors, this step is parallelized via a shared memory scheme with OpenMP.[44] Finally, two-electron integrals are calculated, in parallel, using the resolution of the identity (RI) approximation and are stored in memory for the duration of the calculation.[45]

3.3.3.4 Wave Function Properties

A variety of analyses are implemented for the RAS wave functions, with the first being the spin-expectation value, $\langle S^2 \rangle$. The one-electron reduced density matrix (1eRDM) and the transition density matrix (TDM) are also calculated, from which natural orbitals, dipole moments, and transition moments are obtained. Finally, a restart functionality was developed in which the current guess for each state is saved at the end of each iteration. This is especially useful for long calculations that may be interrupted or for running a newly developed analysis on a converged wave function from a previous calculation.

3.4 Computational Details

RAS-SF was implemented in C++ in a developer's version of the Q-Chem 4.0 software package. All other calculations are run with an unmodified Q-Chem 4.0 install. The frozen core approximation is used for all calculations.

Methylene computations were performed for comparison with a prior FCI study, utilizing the same geometries.[46] The same basis (TZ2P), geometries, and frozen orbitals (one occupied, one virtual) are used. RAS-SF calculations applied a triplet ROHF reference with a (2,2) active space. As TZ2P is a triple-zeta Dunning basis set augmented specifically for the FCI study, there is no corresponding auxiliary basis. Therefore, the large rimp2-cc-pVQZ auxiliary basis was used to ensure accurate computation of the two-electron integrals. Adiabatic corrections to RAS(S,2h,2p)-SF values are minimal and do not change conclusions. See Appendix B.1 for discussion.

Tetramethyleneethane (TME) calculations were set up for comparison to the results of Jordan et al.,[47] and utilize the same geometries. Two sets of RAS-SF calculations were run, one with a triplet reference and (2,2) active space, and one with a septet reference and (6,6) active space. A cc-pVTZ/rimp2-cc-pVTZ basis was used with no frozen virtuals for both sets. As in methylene, adiabatic corrections are not considered for TME. See Appendix B.2 for discussion.

Binuclear transition metal complex calculations utilized the same geometries as in a previous SF study.[28] All calculations utilized Ahlrich's VTZ basis[48] and the corresponding auxiliary basis.[49] Active spaces and references were selected based on the total number of valence electrons on the transition metal centers - leading to a triplet reference with (2,2) active space for Complex 1 and a septet reference with a (6,6) active space for Complexes 2 and 3. No virtuals were frozen.

Tetracene calculations were performed on both a monomer and a dimer. The monomer geometry was optimized with ωB97X-D/cc-pVTZ. Single point energies were obtained with RI-EOM-CCSD and RAS-SF and a cc-pVTZ/rimp2-cc-pVTZ basis. RAS-SF calculations utilized a triplet reference and a (2,2) active space. No virtuals were frozen in any monomer calculation. For the dimer, the one most relevant to singlet fission was chosen, found in the *ab* plane and defined by the [*a b*] translation vector [-½ ½].[50] The dimer geometry was obtained by optimizing the C-H bonds of the crystal structure[51] at the ωB97x-D/cc-pVTZ level. RAS-SF calculations utilized a quintet reference and a (4,4) active space. The basis was chosen to balance cost and accuracy (Table B.5) due to the large system size, settling on cc-pVTZ minus f functions for carbons and cc-pVDZ for hydrogens (C:cc-pVTZ-f/H:cc-pVDZ) along with rimp2-cc-pVTZ. 300 virtuals were frozen, the point at which the energy gaps for the six lowest singlet states were converged to 0.001 eV for RAS(S)-SF calculations (Table B.6). A set of dimer calculations utilizing 6-31G* and 200 frozen virtuals was also done to estimate transition dipole moments.

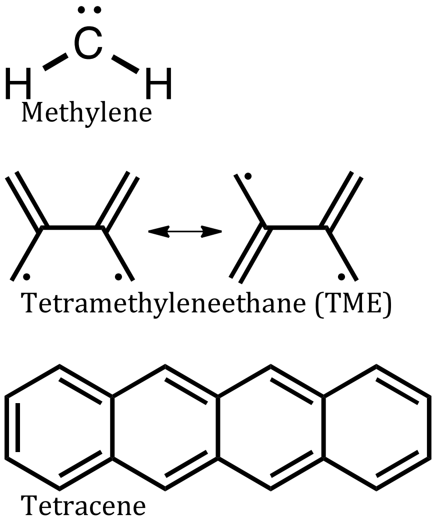


Figure 3.2 Benchmark molecules used in evaluating RAS(S,2h,2p)-SF

3.5 Results and Discussion

3.5.1 Methylene

Methylene is an important test system for excited state electronic structure methods.[52] Its small size enables benchmark-level calculations,[46,53–58] yet its low-lying excited states provide a challenging test case due to varying mixtures of static and dynamic correlation. The four lowest electronic states have been benchmarked at the FCI level with a TZ2P basis by Sherrill et al.[46] The ground ($\tilde{X}$ 1^3B_1) and second excited state ($\tilde{b}$ 1^1B_1) are well described by a single electronic configuration, $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(1b_1)^1$, with different spin couplings. Difficulties arise when tackling the first ($\tilde{a}$ 1^1A_1) and third ($\tilde{c}$ 2^1A_1) excited states, which have significant static correlations due to contributions from $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2$ and $(1a_1)^2(2a_1)^2(1b_2)^2(1b_1)^2$ configurations. A series of RAS-SF computations with minimal and enlarged active spaces are herein compared to the FCI results (Table 3.1). Errors are reported with respect to the FCI energies, with absolute energies provided in the Table B1.

Table 3.1 Errors in RAS-SF methylene calculations compared to the FCI results of Sherrill et al.

	(2,2)			(4,4)
Errors	RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF	RAS(S,2h,2p)-SF
Absolute Energies vs. FCI (mHa)[a]				
$\tilde{X}$	134	117	109	55
$\tilde{a}$	139	117	111	57
$\tilde{b}$	139	124	109	55
$\tilde{c}$	154	153	110	54
Adiabatic Gap vs. FCI (eV)[a]				
$\tilde{X} \rightarrow \tilde{a}$	0.138	0.010	0.063	0.062
$\tilde{X} \rightarrow \tilde{b}$	0.151	0.189	0.002	0.010
$\tilde{X} \rightarrow \tilde{c}$	0.552	0.996	0.024	0.027

[a] FCI results from Sherrill et al[46]

As seen in the upper part of Table 3.1, enlarging the configuration space provides a variational pathway to FCI total energies, but sizable differences in absolute energies are still observed. This is expected given the excitations and active space sizes of the RAS-SF methods, which generate CI spaces much smaller than that of FCI. RAS-SF methods are expected to obtain much lower errors for relative energy levels, i.e. the energy gaps, even with their small configuration spaces. Indeed, adiabatic gap errors are much lower for all SF methods, with RAS(S,2h,2p)-SF performing the best with an order of magnitude reduction in error compared to RAS(h,p)-SF. The effect of enlarging the active space to (4,4) at the RAS(S,2h,2p)-SF level was also investigated, but negligible changes in adiabatic gap errors were found. The triplet reference minimal active space is therefore adequate for accurately capturing the adiabatic energy gaps.

The variance in the $\tilde{X} \rightarrow \tilde{c}$ (i.e. ground to 3rd excited state) gap at different levels of theory deserves further discussion. Table 3.1 shows that RAS(S)-SF almost doubles the error (to roughly 1 eV) with respect to RAS(h,p)-SF. RAS(S,2h,2p)-SF reduces the error to a respectable 0.024 eV. Closer inspection of each excitation levels' contributions to the wave functions sheds some light on this peculiar behavior.

The additional *h-p* excitations of RAS(S)-SF significantly stabilize $\tilde{X}$, contributing 1.4% to its wave function, but only marginally stabilize $\tilde{c}$, contributing only 0.07% to its wave function. The $\tilde{X} \rightarrow \tilde{c}$ gap therefore grows larger in comparison to RAS(h,p)-SF. Moving to RAS(S,2h,2p)-SF, the doubles excitations contribute only 0.42% to $\tilde{X}$, but significantly stabilize $\tilde{c}$ with a contribution of 2.87% to its wave function. Thus, *h-p* excitations stabilize $\tilde{X}$ while doubles excitations stabilize $\tilde{c}$, requiring RAS(S,2h,2p)-SF for a balanced description of the $\tilde{X} \rightarrow \tilde{c}$ gap. This demonstrates that doubles excitations can be vital to the accurate description of gaps between low-lying states, especially when those states have diverse electronic character.

3.5.2 Tetramethyleneethane (TME)

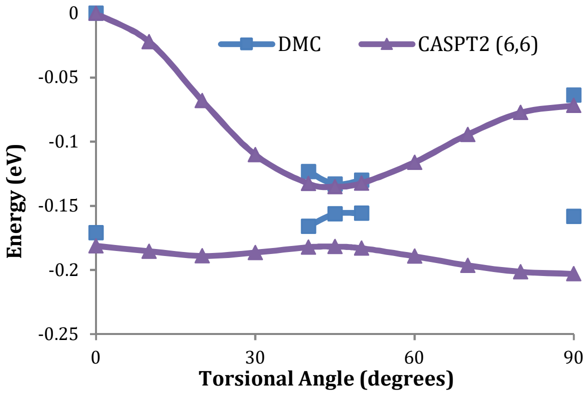
TME is another important benchmark system and the simplest disjoint biradical. Disjoint biradicals exist when the lowest singlet and triplet states are close in energy due to minimal interaction between the two radical electrons. Determination of the ground state of TME is further complicated by the fact that the molecule easily twists around the bridging C-C bond, necessitating scans along this twisting coordinate to determine the lowest energy geometry and state.

A recent study by Jordan et al.[47] utilized diffusion Monte Carlo (DMC) and a truncated cc-pV5Z basis coupled with a pseudopotential to create a torsional potential for TME. This potential scan showed the singlet to be the ground state at all angles, with a slight maximum in the singlet potential at 45°. The study also reported torsional potentials with CASPT2/cc-pVTZ, finding that a (6,6) active space was necessary to generate potentials qualitatively similar to DMC. The DMC and CASPT2(6,6) potentials are reproduced here for comparison (Figure 3.3a). Using these previous findings as a guide, SF calculations with (2,2) and (6,6), active spaces, with triplet and septet references respectively, were run.

The results show that RAS(S,2h,2p)-SF provides a qualitatively improved picture over RAS(h,p)-SF with the minimal (2,2) active space. Specifically, RAS(h,p)-SF has the singlet and triplet surfaces crossing at roughly 45° torsion (Figure 3.3b), which is qualitatively incorrect compared to DMC results (Figure 3.3a). Including higher excitations at either the RAS(S)-SF or RAS(S,2h,2p)-SF levels corrects the qualitative picture (Figure 3.3b), removing the state crossing. These RAS(S,2h,2p)-SF (2,2) potentials qualitatively match those of a prior DDCI study using a CAS(2,2) active space,[59] although a direct comparison cannot be made due to the use of different geometries.

Moving to the enlarged (6,6) active space improves the qualitative RAS(h,p)-SF potential significantly. The incorrect state crossing is removed and it is missing only the singlet maximum at 45°. The effect of the (6,6) space on RAS(S,2h,2p)-SF's potential is more nuanced. The singlet maximum at 45° is recovered, but the triplet surface becomes too high in energy at angles past about 60°. This discrepancy with the benchmark potentials is most likely due to a basis set superposition error, as there is more overlap at the planar geometry versus twisted ones, and the RAS-SF calculations use a triple-zeta basis compared to Jordan's pentuple-zeta basis. Finally, the quantitative errors at critical torsions in Table 3.2 reveal that an enlarged active space noticeably increases the accuracy of RAS(h,p)-SF calculations, while its effect on RAS(S,2h,2p)-SF values is much reduced.

a)



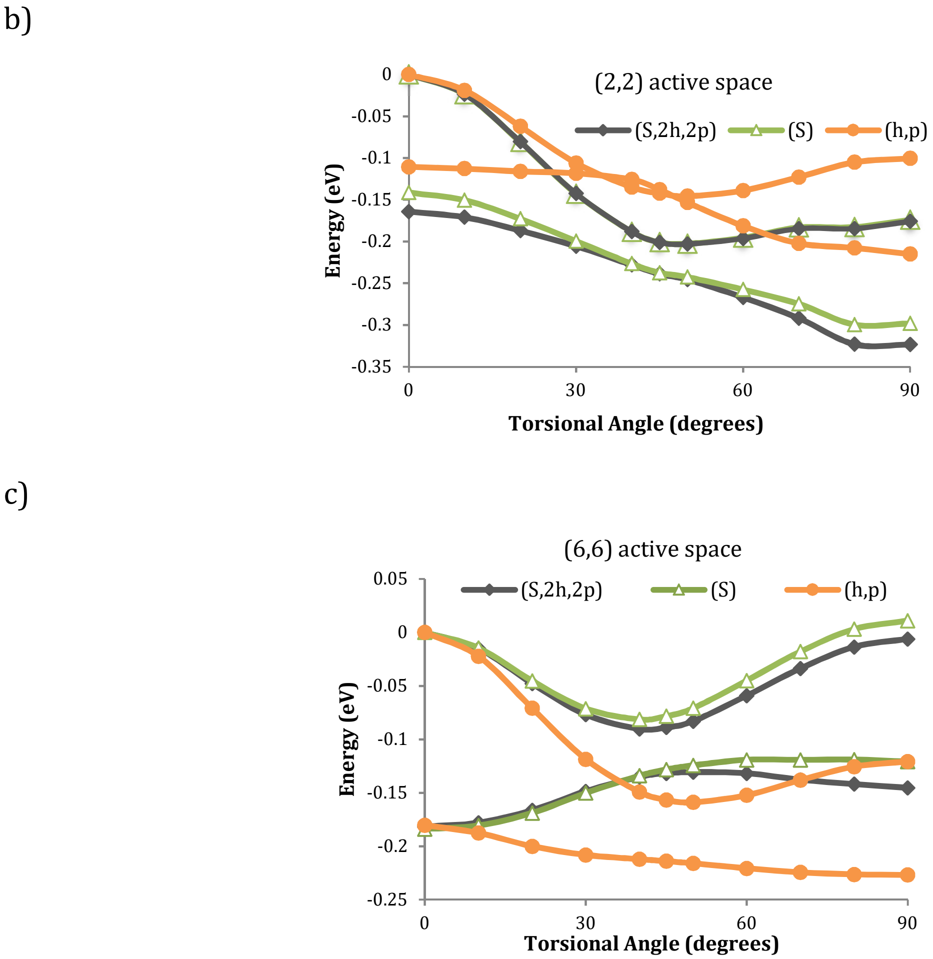


Figure 3.3 Torsional potentials of TME. Each level of theory has two potentials, with the lower one at 0° always belonging to the singlet. All potentials are zeroed to their triplet energy at 0°. a) Torsional potentials from Jordan et al.,[47] created at the DMC and CASPT2(6,6) levels with cc-pV5Z and cc-pVQZ basis sets respectively. b) Torsional potentials from RAS-SF (2,2) calculations. c) Torsional potentials from RAS-SF (6,6) calculations.

Table 3.2 Singlet-triplet energy gaps at critical torsional points of 0, 45 and 90 degrees for tetramethyleneethane

$\Delta E_{S0\rightarrow T1}$ (eV)	DMC[a]	(2,2)			(6,6)		
		RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF	RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF
0°	-0.171	-0.111	-0.141	-0.164	-0.180	-0.184	-0.182
45°	-0.023	0.004	-0.036	-0.038	-0.057	-0.050	-0.043
90°	-0.094	-0.115	-0.122	-0.149	-0.106	-0.132	-0.139

[a] Values from Jordan et al.[47]

The TME results show that RAS(S,2h,2p)-SF improves upon RAS(h,p)-SF, even correcting qualitative inaccuracies when minimal active spaces are used. Furthermore, the negligible improvement in values with enlarged active spaces indicates that RAS(S,2h,2p)-SF can provide very accurate potential energy surfaces even with a minimal active space, which is important for computational tractability in larger systems.

3.5.3 Exchange Coupling in Binuclear Transition Metal Complexes

Single molecule magnetic complexes have interesting applications in spintronics and quantum computing. These are often composed of multiple metal atoms held together by a ligand framework, such that the magnetic properties result from exchange interactions between the neighboring spins of the metal centers. These properties can be modeled by the magnetic exchange coupling, *J*, which can be obtained experimentally or computationally. Computationally, *J* is a parameter in the Heisenberg-Dirac-Van Vleck (HDVV) Hamiltonian,[60] which takes the form of Eq. 3.4. Here, *J* is the exchange coupling between magnetic sites *i* and *j*, and S_i / S_j are spin operators for sites *i* / *j*. As *J* is the only interaction term in the HDVV Hamiltonian, gaps between spin states are dependent only on *J*, which allows one to deduce *J* from electronic structure calculations and the Landé interval rule[61] (Eq. 3.5).

$$\widehat{H}_{HDVV} = -2\sum_{ij} J_{ij}\hat{s}_i\hat{s}_j$$

Equation 3.4 Heisenberg-Dirac-Van Vleck Hamiltonian operator

$$E(S) - E(S-1) = -2SJ$$

Equation 3.5 Magnetic exchange coupling for systems described by HDVV Hamiltonian

Two dichromium complexes, for which existing experimental values were available, have been studied for their magnetic properties.[26] At the RAS(h,p)-SF level of the theory, computed *J*'s were of qualitative accuracy and captured the correct sign, but their magnitudes were much lower than experimental values. Increasing the active space led to minor improvements, and it was posited that inclusion of more dynamic correlation would correct the coupling even further. Indeed, SF-CAS(S)$_n$ calculations by Mayhall et al.[28], where all single excitations are perturbatively included, on these same complexes gave rise to *J* values markedly closer to experimental values (Table 3.3). The inclusion of double's excitations may further increase the accuracy of *J* values determined by SF calculations. The difference in values between perturbative and explicit inclusion of single excitations is also of interest. A divanadium complex was also examined since prior RAS(h,p)-SF calculations gave a qualitatively incorrect *J*.[28]

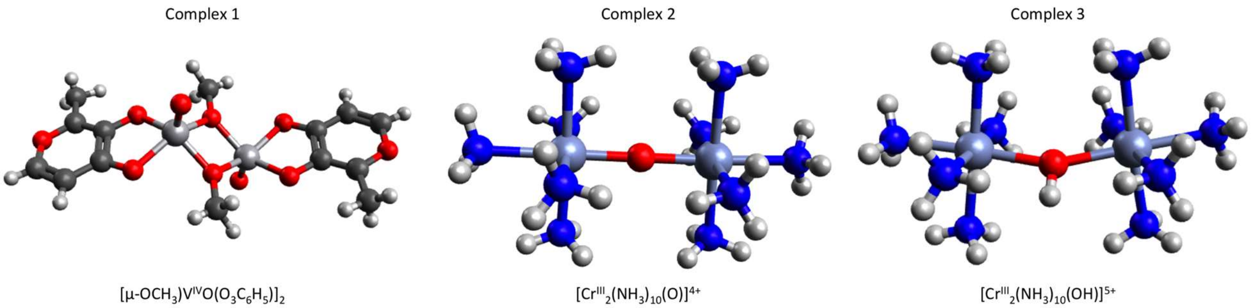


Figure 3.4 Three transition metal complexes for which *J* values were calculated. Complex 1 has one unpaired electron on each V. Since there are two V atoms, SF calculations were run with a (2,2) active space and a triplet reference. Complexes 2 and 3 have three electrons on each Cr, leading to a (6,6)/septet SF calculation.

For RAS(S,2h,2p)-SF calculations, the high-spin references and active spaces were selected based on the total number of valence electrons of the transition metal centers. This lead to assignments of Complex 1: (2,2)/triplet, Complex 2: (6,6)/septet, and Complex 3: (6,6)/septet. The results are compared to experiment as well as the SF-CAS(S)$_n$ calculations of Mayhall et al.[28] in Table 3.3. Ahlrich's VTZ basis[48,49] and the same geometries as in Mayhall et al. were used to allow direct comparison.

Table 3.3 *J*'s (cm-1) determined by SF calculations for a variety of transition metal complexes.

J in cm^{-1}	Complex 1	Complex 2	Complex 3
SF-CAS(S)$_1$[a,b]	-40.9	-150.1	-9.4
SF-CAS(S)[a]	-22.5	-115.8	-7.7
RAS(h,p)-SF[a]	1.1	-63.6	-3.8
RAS(S)-SF	-30.1	-124.7	-8.1
RAS(S,2h,2p)-SF	-33.9	-134.1	-8.2
Exp.	-107[c]	-225[d]	-15.8[e]

[a] from Mayhall et al.[28]
[b] SF-CAS(S)$_1$ values utilize the 500 mH level shift recommended in Mayhall et al.[28]
[c] Reference [93]
[d] Reference [94]
[e] Reference [95]

The results in Table 3.3 indicate that the missing singles excitations (h-p), included either perturbatively or explicitly, correct the J value the most. Although SF-CAS(S)$_1$ performs better than RAS(S)-SF for all complexes, it should be noted that SF-CAS(S)$_1$ results are dependent on an empirical level-shift parameter. SF-CAS(S), utilizing no empirical level-shift, performs slightly worse than RAS(S)-SF for all complexes. Thus, among the methods that include only up to h-p excitations, RAS(S)-SF provides the most accurate J values without the use of empirical parameters. Adding further excitations in the RAS(S,2h,2p)-SF method provides minimal corrections.

Although RAS(S)-SF and RAS(S,2h,2p)-SF J values are markedly improved over RAS(h,p)-SF's with errors ranging from 8 to 100 cm^{-1}, there is still a rather large error relative to experimental J values, from 33 to 50%. This behavior is due to the missing 2h-p and h-2p excitations, which provide a non-negligible contribution to the magnetic exchange value in both antiferromagnetic and ferromagnetic complexes.[62,63] Thus, RAS(S)-SF and RAS(S,2h,2p)-SF are shown to provide J values that are improved over RAS(h,p)-SF but remain only semiquantitative.

3.5.4 Tetracene Singlet Fission

Next-generation organic solar cells may be able to use singlet fission, a multi-exciton generation process,[64,65] to increase their maximum solar conversion efficiency from 33% to 45%.[66,67] RAS(h,p)-SF is well suited to study singlet fission due to its ability to tackle large systems and describe the quadruple radicaloid ^{1}(TT) state that represents singlet recoupling of two triplet excitons. Indeed, RAS(h,p)-SF has been used to highlight the importance of vibronic coupling in pentacene and tetracene singlet fission,[68–70] estimate non-adiabatic couplings,[71–73] and develop models for the dynamics of fission.[74] One possible deficiency of these studies is that RAS(h,p)-SF only qualitatively describes the S_1 and T_1 states. These errors result from RAS(h,p)-SF's limited treatment of dynamic correlation, and are typically corrected for with an energy shift to match experimental values. Below, the effect of correlation on excitonic and multi-excitonic states in RAS-SF computations on the tetracene singlet fission system is investigated and shown to be important for the description of these states.

Table 3.4 Tetracene monomer and dimer calculations. For dimer results, S_1 and S_2 refer to the first and second excited singlets that are not the ^{1}(TT) state.

Tetracene Monomer					
$\Delta E_{S0\rightarrow X}$ (eV)	EOM-CCSD	SOS-CIS(D)	RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF
S_1	3.336	3.086	4.022	3.702	3.506
T_1	1.662	1.726	1.935	1.858	2.033
Tetracene Dimer					
$\Delta E_{S0\rightarrow X}$ (eV)	EOM-CCSD[b]	SOS-CIS(D)	RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF
^{1}(TT)	-	-[a]	3.714	3.821	4.122
S_1	-	2.953	4.038	3.773	3.624
S_2	-	3.064	4.162	3.921	3.761
T_1	-	1.619	1.787	1.861	1.982
^{5}(TT) = Q	-	-[a]	3.741	3.844	4.085

[a] SOS-CIS(D) can only compute single electron excited states.

[b] EOM-CCSD calculations on the tetracene dimer with the selected C:cc-pVTZ-f/H:cc-pVDZ basis are not tractable

3.5.4.1 Tetracene State Energies

The key intermediate of fission, consisting of two triplets coupled into a singlet ^{1}(TT) state, requires a tetracene dimer as a minimal model system. Previous RAS-SF studies have suggested that the energy difference between ^{5}(TT) and ^{1}(TT) can be used to estimate the rate of $^1(TT)\rightarrow 2T_1$ evolution, and thus the overall rate of singlet fission.[71,74,75] This splitting can be quite small, a few tens of meV in tetracene dimers,[74–76] and so the effect of dynamic correlation on this quantity could be significant. To obtain quantitative results, the triple-zeta, polarized, cc-pVTZ basis (minus f functions) for carbon and double-zeta, polarized cc-pVDZ for hydrogen (C:cc-pVTZ-f/H:cc-pVDZ) were chosen based on a series of monomer calculations (Table B.5). These monomer computations showed improved excitation energies for S_1 and minor degradations in the T_1 excitation energies as correlation was added when compared to EOM-CCSD[77] level calculations (Table 3.4). The dimer exciton energies, compared to SOS-CIS(D),[78] similarly improve with increased correlation. SOS-CIS(D) values are used as dimer benchmarks because the large size of the tetracene dimer precludes the use of higher level methods. Neither EOM-CCSD or CIS(D) methods are suitable for describing multi-exciton states and so are not useful for describing the ^{1}(TT) state.

RAS-SF methods are able to describe multi-exciton states, so the computed ^{1}(TT)-^{5}(TT) splitting can be used as an estimate of the ^{1}(TT)-2T$_1$ energy gap (monoexcitonic T$_1$ energies, unfortunately, are semiquantitative (Table 3.4)). This estimate is justified as long as the ^{1}(TT)-^{5}(TT) splitting is accurately calculated and ^{5}(TT) represents independent triplets. The first condition is met because the ROHF quintet reference of RAS-SF provides an excellent description of TT states, and the state similarities allow for significant error cancellations. For the second condition, ^{5}(TT) is a good approximation to 2T$_1$ due to its multi-excitonic nature, which is clearly indicated by unit frontier natural orbital occupations. The use of ^{1}(TT)-^{5}(TT) splitting to approximate the ^{1}(TT)-2T$_1$ energy gap has also been validated by Krylov et al., who demonstrated that a kinetic model utilizing the ^{1}(TT)-^{5}(TT) splitting is able to capture trends of measured singlet fission yields.[74] To illustrate that this estimate remains suitable at the RAS(S,2h,2p)-SF level, as well as to rationalize changes in the ^{1}(TT)-^{5}(TT) splitting, the TT states of the tetracene dimer are examined at various levels of sophistication.

Starting with a separated tetracene dimer, pure multi-exciton states are constructed solely from the doubly-excited configurations of a (4,4) active space. The constraint to doubly-excited configurations prevents interaction with single-exciton states, resulting in pure ^{1}(TT) and ^{5}(TT) wave functions. With separated, non-interacting monomers, pure ^{1}(TT) and ^{5}(TT) are degenerate and are exactly equal in energy to two independent T$_1$ excitons. When the monomers interact in the crystalline dimer geometry, ^{5}(TT) drops below ^{1}(TT) by 0.006 eV (Table 3.5) due to exchange interactions.[64,76]

Table 3.5 Absolute energies (in eV) of pure multi-exciton states at separated and crystalline tetracene dimer geometries. Zero energy is the energy of the degenerate TT states at the separated geometry. Pure multi-exciton states were obtained by restricting the basis to the doubly excited configurations of a (4,4) active space.

Dimer Orientation	Pure ^{1}(TT)	Pure ^{5}(TT)
Separated	0.000	0.000
Crystalline	0.179	0.173

Extending the level of correlation to RAS(h,p)-SF allows the pure ^{1}(TT) to interact with nearby single-exciton states. This mixing leads to the loss of ^{1}(TT)'s pure multi-excitonic character,[71,73] and overall lowers ^{1}(TT)'s total energy. ^{1}(TT) remains predominantly multi-excitonic, however, and is still identifiable via the natural orbital occupations (Table 3.6), which indicate multi-electron excitation character. The ground quintet, ^{5}(TT), described by the ROHF quintet reference, cannot be further improved by *h* or *p* configurations, and thus remains purely multi-excitonic. These two factors result in E(^{1}(TT)) < E(^{5}(TT)) at the RAS(h,p)-SF level (Table 3.4 and Figure 3.5).

Table 3.6 Natural orbital occupancies of the ^{1}(TT) and ^{5}(TT) states of the tetracene dimer. A pure multi-exciton state will have perfectly unit occupancies, as is the case for ^{5}(TT) in RAS(h,p)-SF.

	RAS(h,p)-SF		RAS(S,2h,2p)-SF	
Natural Orbitals	^{1}(TT)	^{5}(TT)	^{1}(TT)	^{5}(TT)
LUNO+1	0.774	1.000	0.420	0.965
LUNO	0.991	1.000	0.948	0.968
HONO	1.009	1.000	1.038	1.032
HONO-1	1.226	1.000	1.593	1.036

Moving to the RAS(S,2h,2p)-SF level incorporates dynamic correlation into TT state descriptions. This results in a more accurate description of ^{1}(TT)'s interaction with nearby singlets, and also correlates ^{5}(TT) past the HF level. At this level, deviations of ^{5}(TT) from a purely double T_1 state may occur due to dynamic correlation and mixing of ^{5}(TT) with higher quintets. The latter, however, is expected to be a small effect since ^{5}(TT) is the ground quintet and is well-separated from excited quintets (the next quintet is 3.0 eV away in the tetracene dimer geometry). Consequently, changes in ^{5}(TT) are due primarily to the addition of dynamic correlation beyond the HF level.

RAS(S,2h,2p)-SF computations noticeably alter the characters of the TT states compared to RAS(h,p)-SF (Table 3.6). ^{1}(TT) shows a further reduction in multi-exciton character, indicating strong mixing of ^{1}(TT) with nearby mono-excited singlet states. On the other hand, ^{5}(TT) natural orbital occupancies change only slightly (less than 0.05 e^-) under the effect of dynamic correlation. Therefore, the ^{1}(TT)-^{5}(TT) approximation remains a useful metric at the RAS(S,2h,2p)-SF level.

Computations at the RAS(h,p)-SF level predict an energy barrier for singlet fission's $^1(TT) \rightarrow 2T_1$ process, placing the ^{1}(TT) state below ^{5}(TT) by 0.1-0.3 eV (Table 3.4 and previous computations[74–76]). The addition of doubles excitations at the RAS(S,2h,2p)-SF level, however, provides more correlation energy to ^{5}(TT) than ^{1}(TT) and reverses the state ordering. With this additional correlation, ^{1}(TT) lies *above* ^{5}(TT) by 0.037 eV (Figure 3.5). When viewing the ^{5}(TT) energy as a proxy for E($2T_1$), this result indicates that the ^{1}(TT) state is unbound, meaning it has a driving force instead of a barrier towards evolution into two independent triplets.

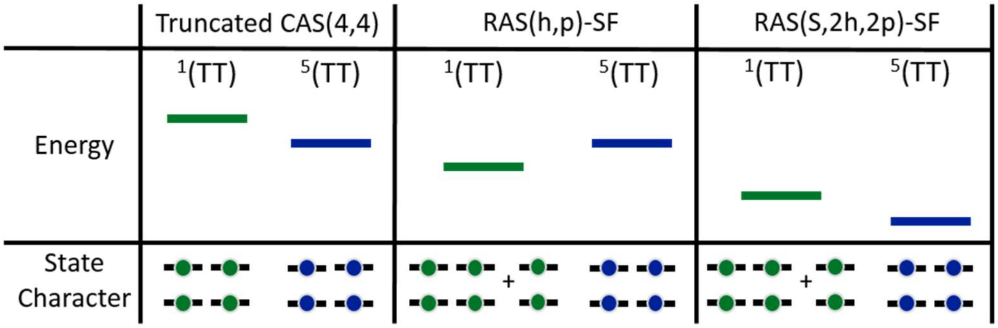


Figure 3.5 ^{1}(TT) and ^{5}(TT) state ordering in crystalline tetracene dimer calculations at varying levels of theory. State characters are represented pictorially by electron occupations of the first four frontier orbitals.

The ^{1}(TT)-^{5}(TT) energy difference also impacts the likelihood of ^{5}(TT) population. Specifically, the small splitting computed in this work and by Krylov et al.[75] raises as a possibility via ^{1}(TT)-^{5}(TT) mixing. Indeed, a recent electron spin resonance (ESR) study unambiguously identified a ^{5}(TT) signal in a covalent dimer of pentacene, where the ^{1}(TT)-^{5}(TT) splitting was estimated to be 80 μeV.[79] The small ^{1}(TT)-^{5}(TT) splitting and downhill location of ^{5}(TT) relative to ^{1}(TT) computed by RAS(S,2h,2p)-SF suggests that ^{5}(TT) will also be populated in tetracene crystals. The possibility of ^{5}(TT) population, and its influence on singlet fission dynamics should therefore be carefully evaluated in future studies.

The above RAS(S,2h,2p)-SF results also offer insight into the ability of tetracene to produce free triplets via singlet fission despite apparently disobeying the energy gap rule, $E(S1) \geq 2E(T1)$.[64,74,80,81] The strong mixing of ^{1}(TT) with nearby singlets means that ^{1}(TT) takes on the delocalized nature of singlet excitons.[69,82] Previous studies showing that entropic considerations make tetracene singlet fission exoergic ($\Delta G < 0$)[74,81] assumed ^{1}(TT) to be localized to neighboring monomers. Replacing this assumption with the more delocalized ^{1}(TT) predicted by RAS(S,2h,2p)-SF results in a greater free energy difference for the $^1(TT) \rightarrow 2T_1$ step, as the entropic contribution to the free energy grows due to an increase in sites on which the two independent triplets may be generated.[74] Furthermore, the picture of a delocalized ^{1}(TT) agrees well with experiments showing cooperative singlet and triplet diffusion in tetracene,[83] which would be expected as ^{1}(TT) and S_1 species rapidly interconvert. The increased free energy difference combined with the unbound nature of the ^{1}(TT) state estimated by the ^{1}(TT)-^{5}(TT) splitting suggests that ^{1}(TT), once populated, will readily evolve into two independent triplets.

3.5.4.2 Dimer Transition Dipole Moments

The ^{1}(TT) spectrum has never before been simulated using electronic structure methods, making it an interesting target for correlated RAS-SF approaches. Experimentally, tetracene solids have been probed with transient absorption (TA) for evidence of the formation of ^{1}(TT) and independent triplets.[84–86] It is generally difficult, however, to directly monitor the ^{1}(TT) state due to overlapping $S_1 \rightarrow S_n$, $T_1 \rightarrow T_n$, ground state bleach, and stimulated emission signals. Though difficult, these assignments have been made for a few systems based on comparisons between short and long time-delayed spectra.[75,87,88] For the present study, transition dipole moments (TDM) have been implemented in the RAS-SF family of methods, enabling examination of the multi-excitonic ^{1}(TT) state spectrum. These excited state absorption signatures are herein compared to S_1 and T_1 TDMs to help identify peaks in the tetracene spectrum. The wide span of transition energies in experimental TA spectrums, 400-800 nm or 1.5-3.0 eV, necessitates calculating a large number of states, so the modest 6-31G* basis is used for RAS(S,2h,2p)-SF TDM computations. Table 3.7 tabulates the strongest transitions, with a more complete tabulation found in the publication corresponding to this Chapter.[89]

Table 3.7 The strongest transitions calculated from the tetracene dimer for the states of interest.

S_1		$^1(TT)$		T_1	
Oscillator Strength	Transition (eV/nm)	Oscillator Strength	Transition (eV/nm)	Oscillator Strength	Transition (eV/nm)
0.07	1.16/1069	0.10	2.62/473	0.08	2.05/605
0.15	2.98/416	0.22	2.86/434	0.07	2.18/569
				0.18	2.96/419

According to RAS(S,2h,2p)-SF, the strongest transitions from S_1, $^1(TT)$ and T_1 take place in the congested 400-550 nm spectral region. This mimics experimental results, where signal overlap in this region has been frequently observed.[84–86] Clearer indications of features correlated with S_1 and T_1 population dynamics have been experimentally identified at longer wavelengths for S_1 (680 nm) and T_1 (~800 nm) states.[83,85,86] These features, however, are broad with low intensities[86] and so it may be difficult to find a single transition that correlates with them. Indeed, while RAS(S,2h,2p)-SF results[89] do show transitions near these lower energies, none particularly stand out in transition strength, making a clear assignment difficult. Overall, RAS(S,2h,2p)-SF qualitatively describes the TA spectra of tetracene when the overall oscillator strengths and absorption wave lengths are considered.

Table 3.8 Components along the crystallographic axes of the TDMs for the transitions probed by polarized transient absorption experiments ($S_0 \rightarrow S_1$, $S_1 \rightarrow S_n$, $T_1 \rightarrow T_2$) and the strongest transitions

	TDM Components (Debye)		
	a	*b*	*c*
$S_0 \rightarrow S_1$ at 3.84 eV	-0.05	-0.44	-0.05
$T_1 \rightarrow T_2$ at 2.05 eV	0.34	-0.35	0.00
$S_1 \rightarrow S_n$ at 2.98 eV	0.21	0.36	-0.27
$^1(TT) \rightarrow S_n$ at 2.62 eV	-0.29	0.40	0.00
$^1(TT) \rightarrow S_n$ at 2.86 eV	-0.42	0.53	-0.04
$T_1 \rightarrow T_n$ at 2.96 eV	-0.42	0.41	0.00

Further insight into the signature of the $^1(TT)$ state can be obtained by examining individual TDM components. In general, there is some dispute as to whether TA signals of the $^1(TT)$ state will resemble T_1 signals or not, with either view having some experimental support.[75,87,88] RAS(S,2h,2p)-SF results (vide supra) suggest that mixing of the pure, multi-excitonic ^{1}(TT) with singly excited configurations could significantly affect the triplet pair structure of the resulting state. Computed TDM orientations (Table 3.8) match well with polarized TA experiments which show that the $S_0 \rightarrow S_1$ TDM is strongly aligned along the *b* axis[85,90] and that $T_1 \rightarrow T_2$ TDMs have a larger component along the *a*-axis than $S_1 \rightarrow S_n$ TDMs.[83,91] More interestingly, the components of the strongest transitions of each state show a strong resemblance between the ^{1}(TT) and T_1 TDMs, which have strong *a* and *b* components and a negligible *c* component, in contrast to $S_1 \rightarrow S_n$'s sizeable *c*-component. These signals cannot be fully separated in polarized TA spectra, however, as all the excited state transitions have considerable *a* and *b* components. The similarity of ^{1}(TT) and T_1 TDMs indicate that TA features corresponding to ^{1}(TT) in tetracene, if observed, should be similar to T_1 absorption features, in agreement with recent TA studies of concentrated TIPS-tetracene solutions[87] and pentacene aggregates.[88]

3.5.5 Timings

One of the main advantages of RAS(h,p)-SF is its relatively low cost, which enables it to tackle large systems such as a tetracene dimer. Therefore, timings of representative calculations are given in Table 3.9 to assess the impact of adding doubles excitations to the RAS-SF family of methods. These timings relate solely to the RAS-SF portion of the computations, and do not include the SCF or integral calculations that are required as a preliminary step for any correlated method.

Table 3.9 Representative timings of RAS-SF calculations.

System	Method	Active Space/ # of States	Determinants	Cores	Wall Time (min)[a]
TME (0° geometry)	RAS(h,p)	(2,2)/4	1 140	4	1
	RAS(S)	(2,2)/4	41 490	4	3
	RAS(S,2h,2p)	(2,2)/4	114 076	4	44
TM Complex 2	RAS(h,p)	(6,6)/3	182 800	20	19
	RAS(S)	(6,6)/3	16 562 800	20	1 301
	RAS(S,2h,2p)	(6,6)/3	39 368 080	20	6 859
Tc Dimer	RAS(h,p)	(4,4)/3	31 620	20	4
	RAS(S)	(4,4)/3	4 943 748	20	849
	RAS(S,2h,2p)	(4,4)/3	12 386 800	20	5 509

[a] Calculations are parallelized over the active space, making larger active spaces more efficient when enough cores are available. This effect is seen in the comparison of TM Complex 2 and Tc Dimer wall times.

Though RAS(S,2h,2p) computational times become relatively long compared to RAS(h,p), no other multi-reference method, such as DMRG or stochastic CASSCF, have yet been applied to systems are large and asymmetric as a tetracene dimer. Although DMRG has been applied to the polyacene series,[92] a tetracene dimer loses the quasi-one-dimensionality found in the linear polyacene monomers, making DMRG not obviously applicable. Alternatively, stochastic CASSCF requires many processors to achieve fast speeds.[7] In contrast, the reported computations utilized modest computational resources, with each calculation run on a single node. Finally, the RAS(S)-SF and RAS(S,2h,2p)-SF methods have yet to be fully optimized, so there is room for further improvement in these timings.

3.6 Conclusion

In this work, RAS(S,2h,2p)-SF, a correlated spin flip approach including up to doubles excitations, was introduced, implemented, and tested. RAS(S)-SF, which added a missing class of singles excitations to RAS(h,p)-SF, was also developed. By comparing these three methods in a series of benchmarks, the effect of adding dynamic correlation to SF methods has been revealed. The variety of systems, including small, strongly correlated systems, larger transition metal complexes, and a model tetracene singlet fission system, indicated that adding correlation beyond the level of singles provides a systematic improvement over previous spin flip methods at modest increased cost.

RAS(S,2h,2p)-SF calculations on small systems with strongly correlated electronic states showed a significant improvement over RAS(h,p)-SF, obtaining quantitative results even with minimal active spaces. Investigations of three binuclear transition metal complexes produced more quantitative spin couplings with increasing correlation, with the majority of the correction obtained at the RAS(S)-SF level. Finally, RAS(S,2h,2p)-SF was applied to tetracene dimers, whose large size makes correlated calculations difficult to achieve with standard methods. RAS(S,2h,2p)-SF significantly corrected the energies of single-exciton states, though quantitative accuracy was not reached with a minimal active space. Overall, the additional electronic configurations included in RAS(S,2h,2p)-SF lead to more quantitative descriptions of a variety of difficult to simulate excited states.

The tetracene dimer calculations additionally provided new insights into this classic singlet fission system. Correlation of the ^{5}(TT) state at the RAS(S,2h,2p)-SF level reversed the typical ^{1}(TT) < ^{5}(TT) state ordering seen in previous RAS(h,p)-SF calculations, resulting in the interpretation of ^{1}(TT) as being relatively mobile, and not easily localized to a tetracene dimer. Furthermore, the first transition dipole moment calculations of a multi-exciton state in a tetracene dimer are reported. The TDMs qualitatively described experimental tetracene transient absorption spectra and predicted a strong similarity between the TDMs of the strongest ^{1}(TT) and T_1 transitions.

The above findings indicate that the addition of dynamic correlation is important to capturing key details that may otherwise be missed in lower levels of theory. Unfortunately, the findings also highlight the cost of including dynamic correlation with an enlarged CI space. Even excluding certain determinants as in DDCI, calculations on the tetracene dimer had to be carefully set up to enable calculations to be run on current computational resources. Instead of including entire classes of determinants, in which there are sure to be determinants that affect the energy insignificantly, one may instead endeavor to enlarge the CI space with only the significant configurations. Strategies for intelligently expanding the CI space and reducing diagonalization costs are explored in the following Chapters.

3.7 References

1. Roos, B. O., Taylor, P. R. & Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **48,** 157–173 (1980).
2. Shepard, R. in *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2, Volume 69* (ed. Lawley, K. P.) 63–200 (John Wiley & Sons, Inc., 1987).
3. Knecht, S. *et al.* New Approaches for *ab initio* Calculations of Molecules with Strong Electron Correlation. *Chim. Int. J. Chem.* **70,** 244–251 (2016).
4. Chan, G. K.-L. & Sharma, S. The Density Matrix Renormalization Group in Quantum Chemistry. *Annu. Rev. Phys. Chem.* **62,** 465–481 (2011).
5. Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131,** 54106 (2009).

6. Booth, G. H. & Alavi, A. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials. *J. Chem. Phys.* **132,** 174104 (2010).
7. Thomas, R. E., Sun, Q., Alavi, A. & Booth, G. H. Stochastic Multiconfigurational Self-Consistent Field Theory. *J. Chem. Theory Comput.* **11,** 5316–5325 (2015).
8. Holmes, A. A., Tubman, N. M. & Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **12,** 3674–3680 (2016).
9. Sharma, S., Holmes, A. A., Jeanmairet, G., Alavi, A. & Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **13,** 1595–1604 (2017).
10. Tubman, N. M., Lee, J., Takeshita, T. Y., Head-Gordon, M. & Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **145,** 44112 (2016).
11. Schriber, J. B. & Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **144,** 161106 (2016).
12. Zhang, T. & Evangelista, F. A. A Deterministic Projector Configuration Interaction Approach for the Ground State of Quantum Many-Body Systems. *J. Chem. Theory Comput.* **12,** 4326–4337 (2016).
13. Krylov, A. I. Size-consistent wave functions for bond-breaking: the equation-of-motion spin-flip model. *Chem. Phys. Lett.* **338,** 375–384 (2001).
14. Krylov, A. I. Spin-flip configuration interaction: an electronic structure model that is both variational and size-consistent. *Chem. Phys. Lett.* **350,** 522–530 (2001).
15. Krylov, A. I., Slipchenko, L. V. & Levchenko, S. V. *Electron Correlation Methodology*. **958,** (American Chemical Society, 2007).
16. Shu, Y., Hohenstein, E. G. & Levine, B. G. Configuration interaction singles natural orbitals: An orbital basis for an efficient and size intensive multireference description of electronic excited states. *J. Chem. Phys.* **142,** 24102 (2015).
17. Shavitt, I. in *Methods of Electronic Structure Theory* 189–275 (Springer US, 1977).
18. Davidson, E. R. & McMurchie, L. E. in *Excited States* **5,** 1–39 (Elsevier, 1982).
19. Rico, R. J. & Head-Gordon, M. Single-reference theories of molecular excited states with single and double substitutions. *Chem. Phys. Lett.* **213,** 224–232 (1993).
20. Sears, J. S., Sherrill, C. D. & Krylov, A. I. A spin-complete version of the spin-flip approach to bond breaking: What is the impact of obtaining spin eigenfunctions? *J. Chem. Phys.* **118,** 9084–9094 (2003).
21. Casanova, D. & Head-Gordon, M. The spin-flip extended single excitation configuration interaction method. *J. Chem. Phys.* **129,** 1–13 (2008).
22. Olsen, J., Roos, B. O., Jørgensen, P. & Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* **89,** 2185 (1988).
23. Malmqvist, P. A., Rendell, A. & Roos, B. O. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *J. Phys. Chem.* **94,** 5477–5482 (1990).
24. Casanova, D. & Head-Gordon, M. Restricted active space spin-flip configuration interaction approach: theory, implementation and examples. *Phys. Chem. Chem. Phys.*

11, 9779–90 (2009).

25. Zimmerman, P. M., Bell, F., Goldey, M., Bell, A. T. & Head-Gordon, M. Restricted active space spin-flip configuration interaction: Theory and examples for multiple spin flips with odd numbers of electrons. *J. Chem. Phys.* **137,** 164110 (2012).
26. Bell, F., Zimmerman, P. M., Casanova, D., Goldey, M. & Head-Gordon, M. Restricted active space spin-flip (RAS-SF) with arbitrary number of spin-flips. *Phys. Chem. Chem. Phys.* **15,** 358–66 (2013).
27. Mayhall, N. J., Goldey, M. & Head-Gordon, M. A quasidegenerate 2nd-order perturbation theory approximation to RAS-nSF for excited states and strong correlations. *J. Chem. Theory Comput.* **10,** 589–599 (2014).
28. Mayhall, N. J. & Head-Gordon, M. Increasing spin-flips and decreasing cost: Perturbative corrections for external singles to the complete active space spin flip model for low-lying excited states and strong correlation. *J. Chem. Phys.* **141,** 44112 (2014).
29. Levchenko, S. V. & Krylov, A. I. Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: Theory and application to cyclobutadiene. *J. Chem. Phys.* **120,** 175–185 (2004).
30. Casanova, D., Slipchenko, L. V., Krylov, A. I. & Head-Gordon, M. Double spin-flip approach within equation-of-motion coupled cluster and configuration interaction formalisms: Theory, implementation, and examples. *J. Chem. Phys.* **130,** 44103 (2009).
31. Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space. *Annu. Rev. Phys. Chem.* **59,** 433–462 (2008).
32. Miralles, J., Castell, O., Caballol, R. & Malrieu, J.-P. Specific CI calculation of energy differences: Transition energies and bond energies. *Chem. Phys.* **172,** 33–43 (1993).
33. García, V. M., Castell, O., Caballol, R. & Malrieu, J. P. An iterative difference-dedicated configuration interaction. Proposal and test studies. *Chem. Phys. Lett.* **238,** 222–229 (1995).
34. García, V. M., Castell, O., Reguero, M. & Caballol, R. Singlet-triplet energy gap in halogen-substituted carbenes and silylenes: a difference-dedicated configuration interaction calculation. *Mol. Phys.* **87,** 1395–1404 (1996).
35. Castell, O., García, V. M., Bo, C. & Caballol, R. Relative stability of the $^3A_2$, $^1A_2$, and $^1A_1$ states of phenylnitrene: A difference-dedicated configuration interaction calculation. *J. Comput. Chem.* **17,** 42–48 (1996).
36. García, V. M., Reguero, M. & Caballol, R. Application of the iterative difference-dedicated configuration interaction method to the determination of excitation energies in some benchmark systems: Be, CH^+ , BH and CH_2. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **98,** 50–56 (1997).
37. Zapata-Rivera, J., Caballol, R. & Calzado, C. J. Electronic structure and relative stability of 1:1 Cu-O_2 adducts from difference-dedicated configuration interaction calculations. *J. Comput. Chem.* **32,** 1144–1158 (2011).
38. Cremer, D. From configuration interaction to coupled cluster theory: The quadratic configuration interaction approach. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **3,** 482–503 (2013).
39. Knowles, P. J. & Handy, N. C. A new determinant-based full configuration interaction

method. *Chem. Phys. Lett.* **111,** 315–321 (1984).

40. Knowles, P. J. & Handy, N. C. A determinant based full configuration interaction program. *Comput. Phys. Commun.* **54,** 75–83 (1989).
41. Szabo, A. & Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. (Dover Publications, 1989).
42. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **17,** 87–94 (1975).
43. Leininger, M. L., Sherrill, C. D., Allen, W. D. & Schaefer, H. F. Systematic Study of Selected Diagonalization Methods for Configuration Interaction Matrices. *J. Comput. Chem.* **22,** 1574–1589 (2001).
44. OpenMP Architecture Review Board, see http://www.openmp.org.
45. Feyereisen, M., Fitzgerald, G. & Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **208,** 359–363 (1993).
46. Sherrill, C. D., Leininger, M. L., Van Huis, T. J. & Schaefer, H. F. Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (TZ2P) basis. *J. Chem. Phys.* **108,** 1040 (1998).
47. Pozun, Z. D., Su, X. & Jordan, K. D. Establishing the ground state of the disjoint diradical tetramethyleneethane with quantum Monte Carlo. *J. Am. Chem. Soc.* **135,** 13862–13869 (2013).
48. Schäfer, A., Horn, H. & Ahlrichs, R. Fully Optimized Contracted Gaussian-Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **97,** 2571–2577 (1992).
49. Weigend, F., Häser, M., Patzelt, H. & Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **294,** 143–152 (1998).
50. Berkelbach, T. C., Hybertsen, M. S. & Reichman, D. R. Microscopic theory of singlet exciton fission. II. Application to pentacene dimers and the role of superexchange. *J. Chem. Phys.* **138,** 114103 (2013).
51. Holmes, D., Kumaraswamy, S., Matzger, A. J. & Vollhardt, K. P. C. On the Nature of Nonplanarity in the [N]Phenylenes. *Chem. - A Eur. J.* **5,** 3399–3412 (1999).
52. SCHAEFER, H. F. Methylene: A Paradigm for Computational Quantum Chemistry. *Science.* **231,** 1100–1107 (1986).
53. Bauschlicher, C. W. & Taylor, P. R. A full CI treatment of the $^1A_1$–$^3B_1$ separation in methylene. *J. Chem. Phys.* **85,** 6510 (1986).
54. Sherrill, C. D., Van Huis, T. J., Yamaguchi, Y. & Schaefer, H. F. Full configuration interaction benchmarks for the states of methylene. *J. Mol. Struct. THEOCHEM* **400,** 139–156 (1997).
55. Reynolds, P. J., Dupuis, M. & Lester, W. A. Quantum Monte Carlo calculation of the singlet–triplet splitting in methylene. *J. Chem. Phys.* **82,** 1983 (1985).
56. Zimmerman, P. M., Toulouse, J., Zhang, Z., Musgrave, C. B. & Umrigar, C. J. Excited states of methylene from quantum Monte Carlo. *J. Chem. Phys.* **131,** 124103 (2009).
57. Włoch, M., Gour, J. R. & Piecuch, P. Extension of the renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed hamiltonian to open-shell systems: A benchmark study. *J. Phys. Chem. A* **111,** 11359–11382 (2007).
58. Slipchenko, L. V. & Krylov, A. I. Singlet-triplet gaps in diradicals by the spin-flip

approach: A benchmark study. *J. Chem. Phys.* **117,** 4694–4708 (2002).
59. Rodriguez, E., Reguero, M. & Caballol, R. The Controversial Ground State of Tetramethyleneethane. An ab Initio CI Study. *J. Phys. Chem. A* **104,** 6253 (2000).
60. van Vleck, J. H. *The Theory of Electric and Magnetic Susceptibilities*. (Clarendon Press, 1932).
61. Landé, A. Termstruktur und Zeemaneffekt der Multipletts. *Zeitschrift für Phys.* **15,** 189–205 (1923).
62. Calzado, C. J., Cabrero, J., Malrieu, J. P. & Caballol, R. Analysis of the magnetic coupling in binuclear complexes. I. Physics of the coupling. *J. Chem. Phys.* **116,** 2728 (2002).
63. Calzado, C. J., Angeli, C., Taratiel, D., Caballol, R. & Malrieu, J.-P. Analysis of the magnetic coupling in binuclear systems. III. The role of the ligand to metal charge transfer excitations revisited. *J. Chem. Phys.* **131,** 44327 (2009).
64. Smith, M. B. & Michl, J. Singlet fission. *Chem. Rev.* **110,** 6891–936 (2010).
65. Smith, M. B. & Michl, J. Recent advances in singlet fission. *Annu. Rev. Phys. Chem.* **64,** 361–86 (2013).
66. Shockley, W. & Queisser, H. J. Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *J. Appl. Phys.* **32,** 510–519 (1961).
67. Tayebjee, M. J. Y., Gray-Weale, A. A. & Schmidt, T. W. Thermodynamic Limit of Exciton Fission Solar Cell Efficiency. *J. Phys. Chem. Lett.* **3,** 2749–2754 (2012).
68. Zimmerman, P. M., Zhang, Z. & Musgrave, C. B. Singlet fission in pentacene through multi-exciton quantum states. *Nat. Chem.* **2,** 648–52 (2010).
69. Zimmerman, P. M., Bell, F., Casanova, D. & Head-Gordon, M. Mechanism for Singlet Fission in Pentacene and Tetracene: From Single Exciton to Two Triplets. *J. Am. Chem. Soc.* **133,** 19944–19952 (2011).
70. Zimmerman, P. M., Musgrave, C. B. & Head-Gordon, M. A Correlated Electron View of Singlet Fission. *Acc. Chem. Res.* **46,** 1339–1347 (2013).
71. Feng, X., Luzanov, A. V. & Krylov, A. I. Fission of Entangled Spins: An Electronic Structure Perspective. *J. Phys. Chem. Lett.* **4,** 3845–3852 (2013).
72. Matsika, S., Feng, X., Luzanov, A. V & Krylov, A. I. What We Can Learn from the Norms of One-Particle Density Matrices, and What We Can't: Some Results for Interstate Properties in Model Singlet Fission Systems. *J. Phys. Chem. A* **118,** 11943–11955 (2014).
73. Luzanov, A. V., Casanova, D., Feng, X. & Krylov, A. I. Quantifying charge resonance and multiexciton character in coupled chromophores by charge and spin cumulant analysis. *J. Chem. Phys.* **142,** 224104 (2015).
74. Kolomeisky, A. B., Feng, X. & Krylov, A. I. A Simple Kinetic Model for Singlet Fission: A Role of Electronic and Entropic Contributions to Macroscopic Rates. *J. Phys. Chem. C* **118,** 5188–5195 (2014).
75. Korovina, N. V. *et al.* Singlet Fission in a Covalently Linked Cofacial Alkynyltetracene Dimer. *J. Am. Chem. Soc.* **138,** 617–627 (2016).
76. Mayhall, N. J. From Model Hamiltonians to ab Initio Hamiltonians and Back Again: Using Single Excitation Quantum Chemistry Methods To Find Multiexciton States in Singlet Fission Materials. *J. Chem. Theory Comput.* **12,** 4263–4273 (2016).
77. Stanton, J. F. & Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **98,** 7029 (1993).

78. Rhee, Y. M. & Head-Gordon, M. Scaled second-order perturbation corrections to configuration interaction singles: efficient and reliable excitation energy methods. *J. Phys. Chem. A* **111,** 5314–26 (2007).
79. Tayebjee, M. J. Y. *et al.* Quintet multiexciton dynamics in singlet fission. *Nat. Phys.* **1,** (2016).
80. Tomkiewicz, Y., Groff, R. P. & Avakian, P. Spectroscopic Approach to Energetics of Exciton Fission and Fusion in Tetracene Crystals. *J. Chem. Phys.* **54,** 4504 (1971).
81. Chan, W., Ligges, M. & Zhu, X. The energy barrier in singlet fission can be overcome through coherent coupling and entropic gain. *Nat. Chem.* **4,** 840–5 (2012).
82. Sharifzadeh, S., Darancet, P., Kronik, L. & Neaton, J. B. Low-Energy Charge-Transfer Excitons in Organic Solids from First-Principles: The Case of Pentacene. *J. Phys. Chem. Lett.* **4,** 2197–2201 (2013).
83. Wan, Y. *et al.* Cooperative singlet and triplet exciton transport in tetracene crystals visualized by ultrafast microscopy. *Nat. Chem.* **7,** 785–792 (2015).
84. Thorsmølle, V. K. *et al.* Morphology effectively controls singlet-triplet exciton relaxation and charge transport in organic semiconductors. *Phys. Rev. Lett.* **102,** 3–6 (2009).
85. Burdett, J. J., Müller, A. M., Gosztola, D. & Bardeen, C. J. Excited state dynamics in solid and monomeric tetracene: The roles of superradiance and exciton fission. *J. Chem. Phys.* **133,** 144506 (2010).
86. Wilson, M. W. B. *et al.* Temperature-Independent Singlet Exciton Fission in Tetracene. *J. Am. Chem. Soc.* **135,** 16680–16688 (2013).
87. Stern, H. L. *et al.* Identification of a triplet pair intermediate in singlet exciton fission in solution. *Proc. Natl. Acad. Sci.* **112,** 7656–7661 (2015).
88. Pensack, R. D. *et al.* Observation of Two Triplet-Pair Intermediates in Singlet Exciton Fission. *J. Phys. Chem. Lett.* **7,** 2370–2375 (2016).
89. Chien, A. D. & Zimmerman, P. M. Recovering dynamic correlation in spin flip configuration interaction through a difference dedicated approach. *J. Chem. Phys.* **146,** 14103 (2017).
90. Lim, S.-H., Bjorklund, T. G., Spano, F. C. & Bardeen, C. J. Exciton Delocalization and Superradiance in Tetracene Thin Films and Nanoaggregates. *Phys. Rev. Lett.* **92,** 107402 (2004).
91. Zhang, B. *et al.* Polarization-dependent exciton dynamics in tetracene single crystals. *J. Chem. Phys.* **141,** 244303 (2014).
92. Hachmann, J., Dorando, J. J., Avilés, M. & Chan, G. K.-L. The radical character of the acenes: A density matrix renormalization group study. *J. Chem. Phys.* **127,** 134309 (2007).
93. Sun, Y. *et al.* [(μ-OCH3)VO(ma)]$_2$, a Strongly Antiferromagnetic Oxovanadium(IV) Dimer. *Inorg. Chem.* **37,** 3119–3121 (1998).
94. Pedersen, E. *et al.* Magnetic Properties and Molecular Structure of the μ-Oxo-bis{pentaamminechromium(III)} Ion. *Acta Chem. Scand.* **26,** 333–342 (1972).
95. Veal, J. T. *et al.* Structural and magnetic characterization of μ-hydroxobis[pentaamminechromium(III)] chloride monohydrate, the acid rhodo chloride complex. *Inorg. Chem.* **12,** 2928–2931 (1973).

Chapter 4: Iterative Submatrix Diagonalization for Large Configuration Interaction Problems

This Chapter largely based upon published work:
Reproduced with permission from Chien, A. D. & Zimmerman, P. M. *Mol. Phys.* **8976,** (2017)

4.1 Abstract

The Davidson method has been highly successful for solving for eigenpairs of the large matrices that are common in quantum chemical simulations. Electronic structure simulations, however, can still easily generate matrices that are too large for current computational resources to handle. Therefore, many strategies have arisen to obtain eigenpairs of sufficient accuracy without considering the full Hamiltonian matrix. This Chapter introduces one such strategy by creating a systematic series of submatrix approximations to the full matrix using natural orbitals. By solving for eigenpairs in this series, the eigenvalue accuracy can be gradually increased until a convergence threshold is reached. Importantly, this allows the series to terminate without ever reaching the full matrix, resulting in lower computational costs and reduced memory demands. Application of the method to the full configuration interaction problem for ground states, excited states, and potential energy scans of various systems shows that the iterative submatrix diagonalization method can systematically control eigenvalue errors and provide substantial cost-savings. This method is therefore expected to be highly useful for large-scale diagonalization problems in electronic structure theory.

4.2 Introduction

One of the main steps in a CI calculation is the diagonalization of a Hamiltonian with dimension the size of the chosen CI space. As seen in Chapter 3, CI spaces in approximate methods can easily reach dimensions of 10^7 or larger, pushing the limits of current computational hardware. Tackling even matrices of this size would be impossible if not for the innovative diagonalization algorithms termed Davidson and Davidson-Liu (hereafter, "Davidson" is used to refer to both algorithms).[1,2,3] By explicitly focusing on the lowest few eigenpairs, the Davidson algorithm enables treatment of relatively large Hamiltonians, especially when combined with use of matrix-vector products, called σ vectors, which can be stored instead of the full Hamiltonian [4-11]. The great success of this algorithm can be traced to its ability to quickly converge when matrices are diagonally dominant, which is a common feature of quantum chemical Hamiltonians. This structural characteristic, however, is not the only feature that may be exploited for systematic, yet rapid convergence. By rearranging the Hamiltonian using transformations that do not affect the eigenvalues, faster convergence and reduced costs may be achieved. One can also focus on reducing the amount of necessary matrix-vector products, called σ vectors, which are the most computationally expensive step of the procedure. Successful elimination of any number of large matrix σ-vector formation steps may significantly reduce the overall cost of Davidson diagonalization. Such strategies will be highly useful in alleviating the rapidly increasing cost of electronic structure simulations with increasing level of theory and system size.

One recent method that utilizes these strategies is incremental full configuration interaction (iFCI), which splits up the FCI problem into *n*-body correlation terms, enabling calculation of FCI quality energies for systems as large as divinylhexatriene.[12,13,14] iFCI is able to tackle such large systems due to convergence of the *n*-body correlation terms at low *n*, which is facilitated by the use of localized occupied orbitals. Even at low *n*, however, there are a significant number of challenging diagonalization problems to be solved. Therefore, cutting costs for each *n*-body Hamiltonian is still important, and this is handled in iFCI by the method described in this article.

In CI electronic structure theories, freezing orbitals is a common cost-saving strategy that places determinants in the Hamiltonian from a subspace of the full molecular orbital set, reducing the effective dimensionality.[15] For instance, the frozen core approximation freezes electrons in core orbitals with respect to the excitation operators of the Hamiltonian. Similarly, a frozen virtual approximation can be used. However, frozen virtuals result in eigenvalue errors that are uncontrolled when canonical molecular orbitals (CMOs) are used, as there is no guarantee that higher-lying orbitals are less important to states of interest. Instead, one can use approximate natural orbitals (NOs), where freezing the low occupancy orbitals results in errors that grow with the magnitude of the occupancy cutoff.[16-27] Freezing NOs with an occupancy cutoff, however, does not necessarily provide systematic control over the eigenvalue errors because the cutoff parameter introduces a geometry-dependent error.[26,27] To obtain robust control over eigenvalue errors with frozen virtuals, one must turn to new strategies.

In this work, we investigate the error control and cost-savings properties of a new method, first appearing in iFCI, for determining CI wave functions to within a specified eigenvalue accuracy. The method exploits the underlying structure of the CI Hamiltonian when expressed in the basis of NOs, which divide up the full matrix into a systematically convergent series of submatrices. The method will therefore be called Iterative Submatrix Diagonalization (ISD), which can also be used alongside perturbation theory[28-31] as ISD+PT. Importantly, the method requires no new parameters, and a standard energy threshold is reached at convergence. Tests on diagonalization of large Hamiltonians from FCI[7-11, 32-35] with more than 10^{26} elements, will serve to compare ISD to the standard Davidson algorithm. These tests will show that ISD is particularly effective at computing accurate ground state energies, potentials along multiple-bond dissociation coordinates, and singlet-triplet gaps at a fraction of the cost of FCI and with error control.

The article is organized as follows: the Background section describes the Davidson algorithm, natural orbital, and perturbation theory procedures in the context of CI. The Theory section introduces the ISD and ISD+PT algorithms, which is benchmarked in the Results section for ground and excited states as well as potential energy scans.

4.3 Background

4.3.1 Davidson Algorithm

In CI computations the Hamiltonian can easily contain 10^7 or more rows and columns, making storage and full diagonalization of this matrix infeasible.[15] The Davidson and Davidson-Liu algorithms, which solve for single and multiple eigenpairs of a Hamiltonian respectively, can be used without explicit construction of the matrix. To see how this works, consider the Hamiltonian expressed in a basis of determinants Φ_p, Eq. 4.1, where the eigenvectors of this matrix are wave functions. The energy is as in Equation 4.3 for any vector of normalized CI coefficients, c. Only one or a few of the lowest eigenvalues are required in typical situations. This eigenvalue structure suggests that matrix-vector products of the form of Eq. 4.4 can be constructed and stored (compactly) as vectors, so E can be rewritten as in Eq. 4.5.

$$H_{pq} = \langle \Phi_p | H | \Phi_q \rangle$$

Equation 4.1 Hamiltonian matrix elements in the basis of determinants

$$\Psi = \sum_p c_p \Phi_p$$

Equation 4.2 Eigenvectors of Hamiltonian matrix

$$E = \sum_{p,q} c_p H_{pq} c_q$$

Equation 4.3 Energy of wave function

$$\sigma_p = \sum_q H_{pq} c_q$$

Equation 4.4 Sigma vector

$$E = \sum_p c_p \sigma_p$$

Equation 4.5 Energy written with sigma vectors

Modern, direct CI implementations of the Davidson algorithm[1,2] take advantage of equations 4.4 and 4.5 to enable diagonalization of large matrices without requiring storage of the full Hamiltonian. In the Davidson procedure, correction vectors to the current best approximate eigenvector, c', are constructed via

$$\delta = -(H - E)^{-1}(\sigma' - Ec')$$

Equation 4.6 Davidson correction vector

which is orthogonalized against existing vectors c^i to form a new vector c^{i+1}. Constructing new σ^i and c^i vectors in each iteration expands the subspace within which eigenpairs are solved for, continuing until the residual, $r = \|\sigma' - Ec'\| < \epsilon$.

Iterative diagonalization is feasible because H is diagonally dominant. Specifically, the correction vector δ can be formed by approximating the matrix inversion $(H - E)^{-1}$ as $(H^d - E)^{-1}$, where H^d are the diagonal elements of the Hamiltonian. Using this substitution, inversion of $(H^d - E)^{-1}$ is trivial and yet still a reasonably accurate estimate of $(H - E)^{-1}$. In this modern, direct strategy for performing Davidson diagonalization, σ vector formation remains the most computationally demanding step. As such, there is continuous interest in speeding up this step in various methods. For instance, recent papers accomplish this for TD-DFT.[5,6]

4.3.2 Initial Vectors

An initial vector c^0 is required for initiation of the Davidson algorithm. While any choice that approximates the true eigenvector will usually result in convergence after several iterations, more accurate initial vectors will speed up the process. For instance, a closed-shell ground state can be approximated as a single Hartree-Fock determinant, which can represent, for example, 80% or more of the exact wave function. Better initial guesses can be constructed by fully diagonalizing a small, explicitly formed submatrix of the full

Hamiltonian.[36] These initial guesses are especially useful when the Hartree-Fock determinant does not dominate the true eigenvector (i.e., in excited states or other radicals).

Complete guess matrix diagonalization is computationally restricted to a dimensionality of 10^3 - 10^4 determinants, which represents only a small fraction of the complete Hamiltonian. An efficient method will be suggested below to create accurate initial vectors for large dimension matrices.

4.3.3 Natural Orbital Configuration Interaction

Because the CI Hamiltonian grows rapidly with excitation level and system size, methods to reduce this dimensionality are worth examining. Neglecting a subset of virtual orbitals reduces the size of the CI by eliminating determinants where those orbitals would be occupied. NOs form an excellent basis for selecting determinants to be removed from the CI space, performing much better than CMOs in this respect.[16,17,18]

NOs can be formed at low cost by diagonalizing the one-electron density matrix from perturbation theory or lower-level CI computations. The NO occupancies are given by the density matrix eigenvalues, which represent the degree to which each NO contributes to the electron density of the approximate wave function. When virtual-space NOs are formed and eliminated, this method is known as the frozen virtual NO (FVNO) approximation. Unfortunately, FVNOs require a cutoff parameter to control which orbitals are frozen, which leads to geometry-dependent variations in eigenvalue errors.[27] This parameter, however, might be systematically selected by the electronic structure algorithm, as will be described shortly. Because CI techniques are invariant to orbital rotations within the virtual set, using NOs or CMOs results in the same eigenvalues when all orbitals are unfrozen.

4.3.4 Configuration Interaction plus Perturbation Theory

An alternative approach to reduce the size of the CI Hamiltonian is to divide the determinant space into primary and secondary regions. For instance, the CIPSI method (configuration interaction by perturbatively selecting iteratively) and its recent reenvisionings use perturbation theory (PT) to select the most important determinants.[37-43] This strategy places the dominant determinants into the Hamiltonian for diagonalization, and then uses PT to correct for the missing, smaller contributions to the eigenvalues. Symmetry-Adapted-Cluster CI (SAC-CI) uses a similar technique to compute excited state energies.[44] This strategy of determinant selection by PT criteria has found its most recent success in heat-bath CI (HCI), which utilizes a lower cost selection criterion while retaining similar accuracies.[45,46] Furthermore, popular methods such as CASPT2 (complete-active-space second order perturbation theory) also divide determinant spaces into primary and secondary regions, where FCI in an active space is chosen as the primary determinant space.[47,48,49] While all such methods cannot be reviewed in this article, CIPSI, SAC-CI, HCI, and CASPT2 are mentioned as representative methods that utilize determinant partitioning schemes to great effect.

In the limit that the primary determinant space dominates the wave function, PT works well to recover the neglected degrees of freedom. Two common PT techniques are the Møller-Plesset (MP)[30] and Epstein-Nesbet (EN)[28,29] corrections. The MP perturbation for a single determinant, utilizes Fock diagonal elements, F_{xx}, of occupied orbitals i, j, unoccupied orbitals a, b, and two-electron integrals to correct for missing correlations. The EN correction is performed based on the CI wave function, and uses the variational CI energy, E, to help estimate the contributions from the neglected determinants, Φ_q. MP PT is applied in the CASPT2 method,[47,48,49] while EN PT is used in CIPSI methods.[37-40] Both corrections are subject to large, uncontrolled errors whenever the dominators are small. These errors can possibly be avoided by ensuring that the most important determinants are in the primary CI space.

$$E^{MP} = -\frac{1}{4}\sum_{i,j,a,b}\frac{\langle ab||ij\rangle^2}{F_{aa}+F_{bb}-F_{ii}-F_{jj}}$$

Equation 4.7 Second-order Møller-Plesset energy correction

$$E^{EN} = -\sum_{q}\frac{\langle\Psi|H|\Phi_q\rangle^2}{H_{qq}-E}$$

Equation 4.8 Second-order Epstein-Nesbet energy correction

EN and MP theories are typically, but not exclusively, used in CI wave functions expressed in a basis of CMOs. In the next section, compact expansions of the CI wave function using NO partitioning will provide an alternative means to select the primary CI determinants and their complementary PT space.

4.4 Theory

4.4.1 Iterative Submatrix Diagonalization Method

Specific considerations from the discussion on CI, NOs, and PT will lead to the proposed method. Because state descriptions are dominated by a minority of determinants,[50] being able to select these prior to the CI computation is desirable and will be a central element in this approach. When using FVNOs the CI space is reduced in dimension and accuracy, but importantly, the accuracy is systematically improvable with decreasing NO occupancy cutoff, η. PT corrections to CI are expected to perform well for small η, as low-occupancy NOs result in small contributions to the (truncated) CI wave function.

These ideas now allow us to formulate a systematic series of approximations to a given CI level of theory. While FCI will be used as a prototypical example, being exact in a given orbital basis, this strategy is trivially applicable to lower excitation level CI. Figure 4.1 outlines the overall procedure.

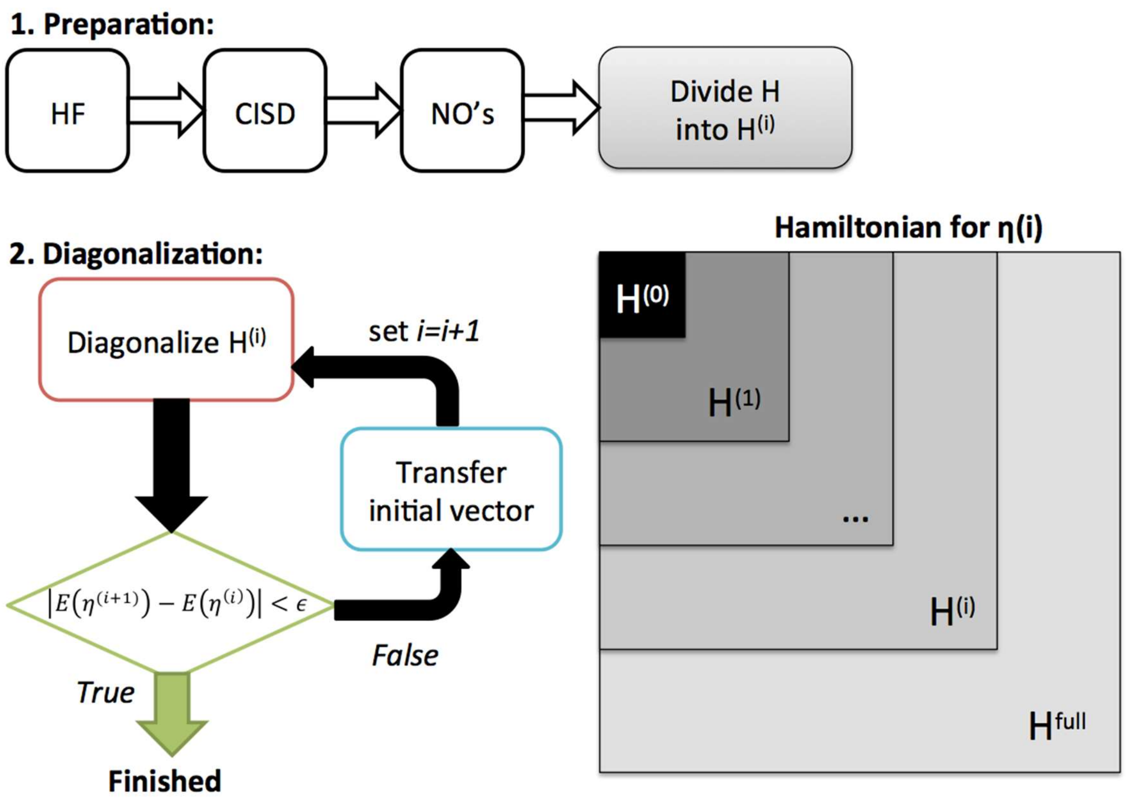


Figure 4.1 Outline of Iterative Submatrix Diagonalization (ISD) algorithm

First, natural orbitals are formed at a low level of theory, CI with singles and doubles substitutions (CISD) from the Hartree-Fock initial determinant (this choice is not unique, and could be substituted by another level of theory).[26,27] After NO formation, occupancy cutoffs, $\eta^{(i)}$, are applied to partition the Hamiltonian into submatrices, $H^{(i)}$, containing only the active NOs for each level of cutoff (see right side of Figure 4.1). At the first iteration, the Davidson algorithm is applied to solve for the eigenpairs of the FVNO FCI submatrix for cutoff $\eta^{(0)}$, denoted $H^{(0)}$. At subsequent iterations, the submatrix of $\mathrm{H}^{\mathrm{full}}$ matrix is enlarged such that $\eta^{(i+1)} < \eta^{(i)}$ and the Davidson algorithm is repeated. The eigenpairs of each submatrix therefore represent an approximate solution to the untruncated CI wave function, which is systematically improved by increasing i.

In this formulation of the CI problem, the exact eigenpairs are recovered as $\eta^{(i)} \to 0$. Exact CI eigenvalues are not typically required, however, but only eigenvalues accurate to within a specified threshold, ϵ (e.g. 1.6 mHa, which is 1 kcal/mol). This allows $\eta^{(i)}$ to decrease with iteration i steps until $|E(\eta^{(i+1)}) - E(\eta^{(i)})| < \epsilon$. The Davidson algorithm, in any case, would only be converged to accuracy ϵ, so no new parameter is introduced to reach the desired convergence threshold. This fact distinguishes this method from typical FVNO implementations, all of which require specification of a cutoff parameter.

$$E\left(\eta^{(i+1)}\right) < E\left(\eta^{(i)}\right)$$

Equation 4.9 Desired ISD property 1

$$\left|\frac{E\left(\eta^{(i+2)}\right) - E\left(\eta^{(i+1)}\right)}{E(\eta^{(i+1)}) - E(\eta^{(i)})}\right| < 1$$

Equation 4.10 Desired ISD property 2

Due to this convergence strategy, the proposed method is most useful when Eq. 4.9 and 4.10 are satisfied, which will result in smooth convergence over the series *i*, *i+1*, *i+2*… when $\eta^{(i)}$ is chosen to be monotonically decreasing. The NO ordering of orbitals and their occupancies is expected to induce property 4.10 onto the CI. This property, however, will be checked numerically in the results section. Due to the variational nature of CI, equation 4.9 is guaranteed because the Hamiltonian dimension increases with decreasing $\eta^{(i)}$.

Since diagonalization is performed on increasing size submatrices of the same full Hamiltonian, each submatrix Davidson procedure can be initiated with the eigenvector of the previous iteration. Therefore, highly accurate guess vectors are available at every iteration, even for large (10^8 or greater) dimension submatrices. Overall, this strategy reduces the number of required Davidson iterations for diagonalizing each submatrix. Importantly, the expensive H^{full} diagonalization step should take fewer iterations.

The new method for finding CI eigenvalues will be denoted the Iterative Submatrix Diagonalization (ISD). One extension of ISD is to perform PT to incorporate effects of the neglected determinants at each iteration.[28,29,30] ISD+PT therefore adds PT corrections at each threshold $\eta^{(i)}$, and could converge at smaller *i* than ISD alone. In the current implementation, the final CI total energy is corrected by doubles excitations out of the primary space via EN PT as described in equation 4.8. While variationality of the wave function is formally lost in ISD+PT, the PT corrections can be small due to the NO partitioning of the CI space, and the energies therefore nearly variational for tight cutoffs, $\eta^{(i)} \ll 1$.

4.4.2 Properties of the Submatrix Davidson Method

Two distinguishing features of ISD and ISD+PT separate it from other methods in this area. First, the CI space is selected via NOs, which inherently capture (and order) the most important orbitals in the wave function. In contrast to FVNO CI and CC,[22-27] which require input of a cutoff parameter (which introduces additional errors), ISD scans over the cutoff value to reach convergence to a standard energy threshold. This scanning strategy also leads to an internal metric that gives control over eigenvalue errors. Methods such as CIPSI[37-40] and CASPT2[47,48,49] select determinants via PT or an active space, and therefore have fundamentally different modes of capturing the key CI determinants. Second, the NO selection of virtual orbitals gives a useful ordering of the CI space, which can be diagonalized in a doubly iterative fashion. By iterating over submatrices of the full problem, each can be diagonalized by the iterative Davidson procedure.

Because ISD and ISD+PT are converged to an energy threshold, ϵ, no new parameters are required. This property is different than methods like CIPSI, which require explicit additional parameters (or implicit parameters, like active space choice in CASPT2). Due to this structure, as $\epsilon \to 0$, ISD and ISD+PT result in the exact eigenpairs of the full Hamiltonian. When ϵ is relatively large, diagonalization of the largest matrix (or matrices) may be avoided entirely due to convergence being reached early. In the latter limit, ISD will have low computational cost.

Finally, ISD is deterministic, variational, and systematically improvable. While ISD+PT loses variationality, it is still improvable by tightening the energy convergence threshold, ϵ. ISD and ISD+PT are applicable to any excitation level of CI,[15] and any diagonalization problem where the Hamiltonian can be written as a function of NOs or NO-like quantities. Importantly, any Davidson-driven CI code capable of handling FVNOs can be modified to perform ISD, without extensive rewriting.

4.5 Computational Details

ISD, ISD+PT, and HCI codes were implemented in C++ in a development version of the Q-Chem 4.3 software package.[51] The CI step of ISD uses a determinant-based, α and β string indexed FCI where c and σ vectors are stored in memory. For all ISD computations, the correlation consistent, polarized cc-pVXZ basis sets were used, where X = D, T.[52] Davidson diagonalization was considered converged when the residual is less than 0.000002 Ha (errors of ~1 μHa are expected). Two electron integrals are computed using the resolution-of-the-identity approximation and the RIMP2-cc-pVTZ basis set.[53] Point group symmetry is not invoked in any ISD or ISD+PT computation. ORCA 4.0[54] was used to perform CIPSI computations via ICE-CI and obtain timings on the OH cation, with energies converged to 1 mHa. The geometry of OH was optimized using closed-shell MP2 with the cc-pVQZ basis set (1.0226 Å). The FCI-quality geometries of methylene were taken from Sherrill et al.[55]

4.6 Results and Discussion

4.6.1 Convergence of ISD and ISD+PT

To show how ISD and ISD+PT behave in practice, the convergence of these methods is now benchmarked. Specifically, the errors in ISD total energies of the ground state, triplet OH cation in the cc-pVDZ and cc-pVTZ basis sets will be compared to FCI with a range of increasingly tight NO cutoffs. This test case contains 6 active and 2 frozen core electrons, making the dimension of the FCI space 665,856 determinants for cc-pVDZ, and 152 million for cc-pVTZ. Figures 4.2 and 4.3 show the errors in energy for CI and CI+PT using the ISD partitioning of the Hamiltonian. The sequence of cutoff values was selected as

$$\eta^{(i)} = 10^{-(5+i)/2}$$

such that the first NO cutoff at $i = 1$ is 10^{-3}=0.001, the second at $i = 2$ is $10^{-3.5}$=0.00032, and so on until all NOs are included. This sequence was chosen because at η =10^{-3}, 13 of 43 orbitals are active in the CI, but only 5 orbitals would be active at η =10^{-2}. Therefore, the cutoff at iteration 1 was sufficiently loose to produce a computationally inexpensive, but nontrivial starting point for the series of submatrix computations.

As the submatrix expands with each iteration of ISD the change in energy compared to the previous iteration, Eq. 4.11, can provide an estimate of the total error, Eq. 4.12. Importantly, if $\Delta\mathrm{E}^{(i)}$ is used as a metric for convergence, we require Eq. 4.13 be obeyed. Figure 4.2 shows that Eq. 4.13 holds for the OH cation, and later results show that Eq. 4.13 holds for all other presented test cases as well. This is also demonstrated in Table 4.1, where the ratio of the total error to ΔE is less than ½ at iteration 3 and later, for both DZ and TZ basis sets. The same convergence metric holds for ISD+PT (Figure 4.3).

$$\Delta\mathrm{E}^{(i)} = \left|E\left(\eta^{(i)}\right) - E\left(\eta^{(i-1)}\right)\right|$$

Equation 4.11 Change in energy between ISD iterations

$$\left|E\left(\eta^{(\infty)}\right) - E\left(\eta^{(i)}\right)\right|$$

Equation 4.12 Total error within an ISD iteration

$$\Delta\mathrm{E}^{(i)} > \left|\mathrm{E}\left(\eta^{(\infty)}\right) - E\left(\eta^{(i)}\right)\right|$$

Equation 4.13 Desired ISD property 3

Equation 4.13 is not guaranteed to be satisfied, and would be difficult to check for larger systems where FCI-level results cannot be obtained. Depending on the NO occupancy distribution, it is possible that Eq. 4.13 could be violated and, if $\Delta\mathrm{E}^{(i)}$ is small enough, lead to premature convergence of ISD. In practice, however, we have not seen this occur when good NOs are available. The sequence of $\eta^{(i)}$ values is fixed and has sufficiently large gaps (i.e. half-orders of magnitude) such that $\Delta\mathrm{E}^{(i)}$ tend to decrease monotonically, and thus the procedure is expected to be transferable.

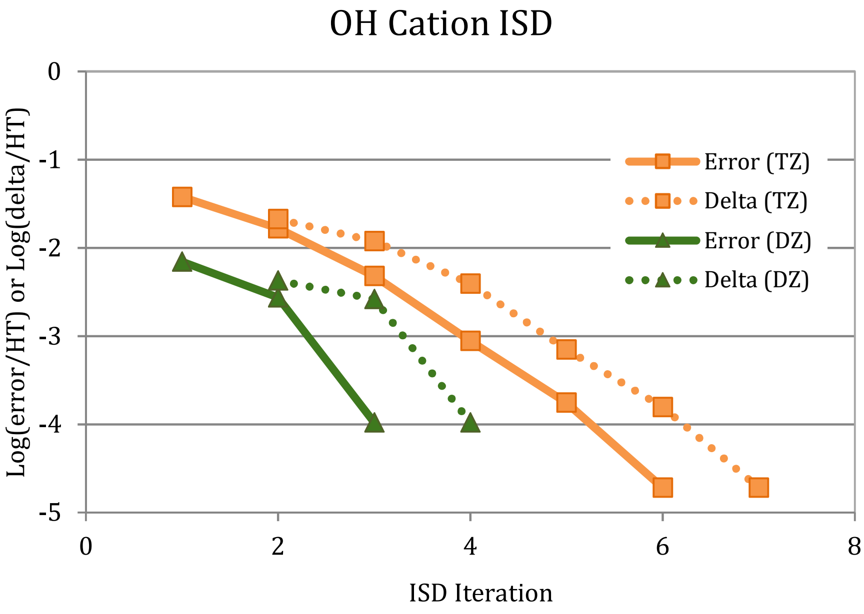


Figure 4.2 Convergence of the ISD energy as a function of iteration number for the OH cation in its ground triplet state. Delta is the absolute change in energy from iteration *i-1* to *i*. For *i* = 4 (DZ) and *i* = 7 (TZ), the errors are zero and not shown.

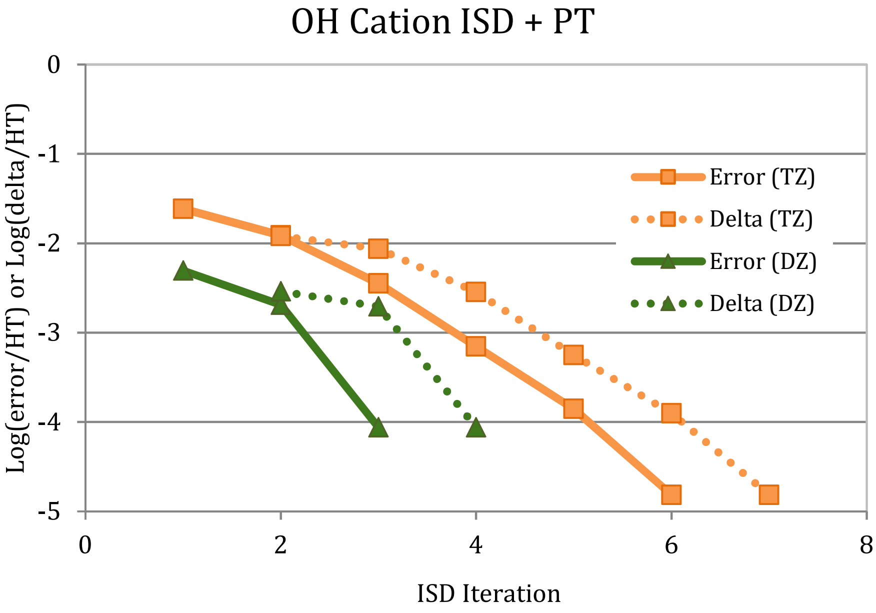


Figure 4.3 Convergence of the ISD+PT energy as a function of iteration number for the OH cation in its ground triplet state. Delta is the absolute change in energy from iteration *i-1* to *i*. For *i* = 4 (DZ) and *i* = 7 (TZ), the errors are zero and not shown.

Table 4.1 Convergence of CI energies as a function of ISD iteration. Energies and errors are in units of Ha.

cc-pVDZ					
Iteration	-log η	E(CI)	Error	$\Delta E^{(i)}$	$\lvert \text{Error}/\Delta E \rvert$
1	3	-75.10352	0.00699		
2	3.5	-75.10778	0.00274	-0.00426	0.64
3	4	-75.11041	0.00010	-0.00263	0.04
4	5.5(∞)	-75.11052	0.00000	-0.00010	0.00
cc-pVTZ					
Iteration	-log η	E(CI)	Error	$\Delta E^{(i)}$	$\lvert \text{Error}/\Delta E \rvert$
1	3	-75.13170	0.03782		
2	3.5	-75.15282	0.01669	-0.02113	0.79
3	4	-75.16471	0.00480	-0.01189	0.40
4	4.5	-75.16864	0.00088	-0.00392	0.22
5	5	-75.16934	0.00018	-0.00070	0.25
6	5.5	-75.16950	0.00002	-0.00016	0.12
7	6(∞)	-75.16951	0.00000	-0.00002	0.00

The results of Figure 4.2, 4.3, and Table 4.1 suggest that ISD provides a well-behaved series of submatrix approximations to the total energy. Importantly, ISD at iteration 5 for cc-pVTZ (iteration 3 for cc-pVDZ) is converged to within 0.7 mHa, or 0.4 kcal/mol, according to the $\Delta E^{(i)}$ metric. This means iterations 6 and 7 do not need to be performed to reach chemical accuracy, saving considerable computational cost. In this case, iteration 6 requires the diagonalization of a matrix with 132 million determinants, and iteration 7 (which is equivalent to FCI) requires diagonalization of a 152 determinant expansion. In contrast, iteration 5's submatrix, $H^{(5)}$, has 71 million determinants, less than ½ that of the FCI dimension.

ISD+PT's performance is improved compared to that of ISD without PT. For the triplet OH cation, the inclusion of PT corrections results in smaller errors and maintains good convergence of the energy. ISD+PT is, however, not necessarily variational, so the remainder of this article will concentrate on standard ISD.

4.6.2 Computational Cost of ISD Convergence

ISD provides a means to converge the eigenvalue to within a specified energy tolerance, ϵ, using submatrix representations of the full Hamiltonian. This property leads to reduced computational effort compared to the standard direct CI Davidson algorithm. This subsection examines the number of Davidson iterations required to diagonalize a given submatrix in order to provide a useful metric for overall computational effort. These iterations are, of course, dominated in cost by the σ-vector generation step, which may also be made faster through efficient algorithms. This aspect is not considered here, and new σ-vector algorithms are discussed elsewhere.[11]

To provide a measure of computational effort, Table 4.2 shows the wall times and number of Davidson iterations required for diagonalization of the submatrices and full matrix Hamiltonian for the OH cation triplet in the cc-pVTZ basis. For the smaller matrices, ISD requires approximately the same number of iterations compared to standard Davidson. For larger matrices, in this case $i>3$, ISD requires fewer Davidson iterations to converge.

The enhanced convergence for submatrix $H^{(i)}$ is due to the high-quality initial vectors provided by the $H^{(i-1)}$ diagonalization. This point is reflected by the initial residuals for the larger matrices. For instance, at $i = 6$, ISD has an initial residual of 0.0014 Ha, while standard Davidson starts at 0.6980 and reaches a residual of less than 0.0014 only after 5 iterations. Therefore, the large matrix initial guess of ISD is an orders-of-magnitude better approximation to the true eigenvalue. Overall, the reduced number of Davidson iterations required for large matrix diagonalization reduces the late-stage computational effort of ISD as well as the memory requirements for storage of the c and σ vectors. For full matrix diagonalization, preliminary tests indicate that the ISD total computational time (including all iterations) is approximately the same as standard Davidson for the full matrix, but this may be improvable in the future. Despite this indication, ISD does provide a significant improvement in computational cost over standard direct CI Davidson whenever it converges before reaching the final Hamiltonian iteration.

Table 4.2 Dimensions of CI and cost comparison between standard and ISD Davidson algorithms for OH cation in the cc-pVTZ basis.

ISD Iteration	Frozen Virtuals	Active Orbitals	Dimension of CI	Davidson Diagonalization Iterations		Davidson Wall Times (sec)[a]		ISD Error (Ha)
				Standard	ISD	Standard	ISD[b]	
1	30	13	81,796	6	6	2.9	2.2	0.037819
2	23	20	1,299,600	7	7	20.5	13.5	0.016691
3	16	27	8,555,625	8	7	174.3	155.3	0.004804
4	10	33	29,767,936	9	7	1198.4	912.5	0.000880
5	5	38	71,166,096	10	6	4209.3	2772.2	0.000176
6	1	42	131,790,400	11	5	11172.1	5485.0	0.000019
7	0	43	152,300,281	11	4	16355.8	6723.6	0.000000

[a] All wall times come from calculations utilizing 28 cores.
[b] ISD timings are per iteration, not cumulative.

When the eigenvalue convergence tolerance, ϵ, has a moderate value, ISD terminates before reaching the full matrix diagonalization step. As noted above, in the case that $\epsilon = 1.6$ mHa, or chemical accuracy, the computational benefits of ISD become apparent. For the OH cation triplet in the cc-pVTZ basis, the largest CI matrix that needs to be diagonalized has dimension of only 71 million, and requires only 6 Davidson iterations at this step to converge (compared to 11 iterations with matrix dimension 152 million for standard Davidson). This ISD result therefore represents a substantial computational savings over standard diagonalization of the full matrix. Comparison with other methods that systematically cut down the size of the Hamiltonian, such as CIPSI[37-40] and HCI,[45,46] on the OH cation, cc-PVTZ basis (Table C.1) show that ISD performs better than CIPSI, but worse than HCI.

4.6.3 Error Control for ISD

ISD can also give estimates of energy profiles with controlled error. The example system is chosen to be diatomic nitrogen dissociation, which is a common test problem for methods that handle combinations of dynamic and static correlation. The cc-pVTZ basis with 2 frozen core orbitals was used, resulting in a FCI space of 2.1x10^{13} determinants. To obtain FCI-level energies with this determinant space, HCI[45, 46] was used as a CI solver in place of conventional FCI (with ε_1 = 0.5 mHa and ε_2 = 25 μHa). With this large determinant space, $\eta^{(i)} = 10^{-(6+i)/2}$ was found to be most appropriate because convergence at 10^{-3} was quite poor. The N_2 dissociation profile shown in Figure 4.4 compares ISD iterations 2 and 3 for interatomic distances of 0.9 to 2.6 Å in steps of 0.1 Å.

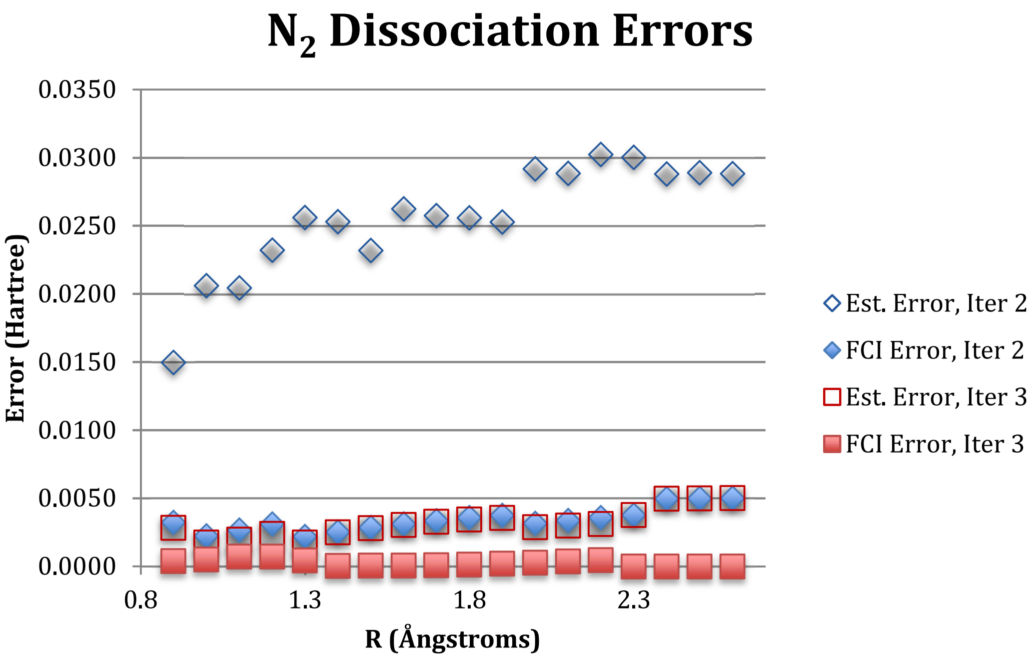


Figure 4.4 Error with respect to FCI-level calculations in diatomic nitrogen dissociation curves with the cc-pVTZ basis for ISD iterations 2 and 3 are shown. Open data points are the estimated errors, while closed one are true errors with respect to FCI-quality energies.

In Figure 4.4, the estimated errors and the true errors with respect to FCI-level energies are shown for ISD at $i = 2$ and $i = 3$, corresponding to FVNO cutoffs $\eta^{(2)} = 10^{-4.0}$ and $\eta^{(3)} = 10^{-4.5}$. As in the OH cation, the errors demonstrate good behavior, monotonically decreasing as η decreases and obeying Eq. 4.13. At $i = 3$, only 38% of the FCI space of determinants are present in $H^{(3)}$ at 1.1 Å. The error estimate, $\Delta E^{(3)}$, is less than 5 mHa for all interatomic distances (the error bars are all negative because ISD is variational), and the energy surface is smooth and asymptotically correct through the range of interatomic distances, as supported by the non-parallelism errors of 2.8 and 0.7

mHa for iterations 2 and 3 respectively. For $i = 2$, however, significant errors are estimated. $\Delta E^{(2)}$ ranges from 15 to 30 mHa, generally increasing as one moves away from the equilibrium well. $\Delta E^{(i)}$ therefore can differentiate between regions of the potential scan that are accurate, and which are most in need of additional convergence. In standard FVNO approximations a constant NO cutoff would result in errors that vary with geometry, making the error estimate of ISD particularly useful for potential energy scans. With ISD, i can be iterated to $i + 1$ whenever necessary to systematically decrease the error and its estimate, ΔE.

4.6.4 Singlet-Triplet Gap of Methylene via ISD

The ISD method can be used for any electronic states where a reasonable set of NOs is available. Interestingly, ISD may also be able to exploit differential correlation, where cancellation of systematic errors in two different states may result in high-quality relative energies. ISD is thus tested here for the CH_2 triplet-to-singlet adiabatic transition, which are the lowest two electronic states of this fragment. Two state-specific computations are performed, one using CISD NOs for the singlet, and the other with CISD NOs for the triplet. The series $\eta^{(i)} = 10^{-(5+i)/2}$ from the OH cation benchmarks was also used here. Methylene’s six valence electrons were correlated, its 1s core orbital kept frozen, and the cc-pVTZ basis was utilized.

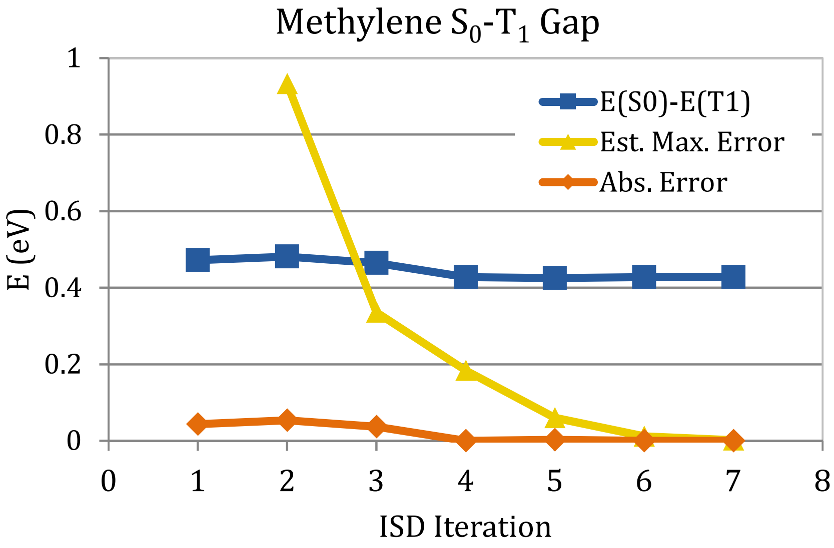


Figure 4.5 Convergence of the triplet-singlet gap in CH_2 as a function of ISD iteration. The estimated maximum error (Est. Max. Error) is $\Delta E^{(i)}(S) + \Delta E^{(i)}(T)$.

Figure 4.5 shows the adiabatic energy gap in methylene as a function of ISD iteration. Excitingly, all ISD steps provide gaps within 0.1 eV of the benchmark value (i.e. FCI for the cc-pVTZ basis). At iterations 4 through 6, ISD is within 0.01 eV of the exact gap, which is 0.430 eV in this basis set (FCI/cc-pVTZ therefore comes near the experimental value is 0.406 eV.[52,56] The actual errors can be compared to the error from the ISD convergence metric, $\Delta E^{(i)}$. Specifically, $\Delta E^{(i)}(S) + \Delta E^{(i)}(T)$ provides an estimate of how far the sum of the two ISD energies are from convergence as a function of iteration.

The estimated maximum error is plotted in Figure 4.5, and is much larger than the realized error. Such a rapid convergence of the realized error can only occur if the correlation contributions from low-occupancy virtual NOs in the two states are similar. In other words, the neglected correlation energy for each cutoff $\eta^{(i)}$, while it may be substantial, is similar in magnitude for the two states. These contributions tend to cancel, leading to fast convergence of the ISD excitation energy with respect to iteration number. As in the previous examples, $\Delta E^{(i)}$ decreases monotonically with decreasing η, demonstrating a useful degree of error control for relative energies. The FVNO cutoffs, $\eta^{(i)}$, are therefore seen to be systematic means of incrementally adding correlating orbitals for the two electronic states.

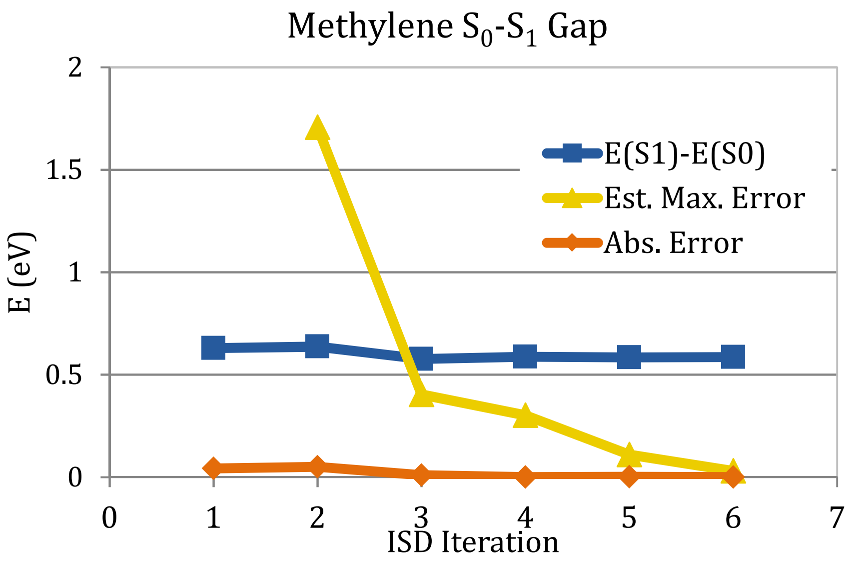
These results suggest that ISD can be used to rapidly compute relative energies to high accuracy. In the case of methylene, the first three iterations' submatrices, $H^{(i)}$ $(i = 1,2,3)$, contain order 10^5, 10^6, and 10^7 determinants, yet provide relative energies within 0.1 eV of FCI which has nearly 10^9 determinants (more precisely, $8.56\text{x}10^8$ determinants). This means 0.1 eV accuracy is reached at less than 1% of the FCI computational cost. Furthermore, $i = 4$ is within 0.01 eV of FCI, yet requires less than 20% of FCI's dimension, around $1.5 \text{ x } 10^8$ determinants. In summary, a converged FCI energy of the triplet-singlet gap in methylene can be computed with dramatically reduced effort using ISD.

4.6.5 Multistate ISD Methylene Calculations

So far, ISD calculations focused on ground state computations with particular spin symmetries, but the description of multiple excited states is also pertinent to many chemical processes. Therefore, a series of multistate ISD calculations were applied to the vertical spectrum of methylene's three lowest singlet states at the ground state (triplet) geometry. These multistate calculations utilize a set of average virtual NOs instead of the state-specific virtual NOs used in previous ISD calculations. The same truncation scheme, frozen core orbitals, and basis set as in the previous, state-specific methylene calculations are used.

Figure 4.6 plots the results of multistate ISD calculations for the S_0-S_2 gaps, also presenting the same error metrics as in Figure 4.5. ISD provides accurate relative energies, with all iterations obtaining vertical gaps within 0.1 eV of FCI/cc-pVTZ benchmark values, and gaps within 0.01 eV being recovered by iteration 3. Estimated errors are again well-behaved with respect to η, and comparisons of the realized error versus the estimated error again show that accurate gaps can be obtained, even at iterations where the neglected correlation energy is substantial. These calculations provide evidence that the use of state-averaged virtual NOs in ISD's truncation scheme is a viable method for accomplishing multistate ISD calculations.

a)



b)

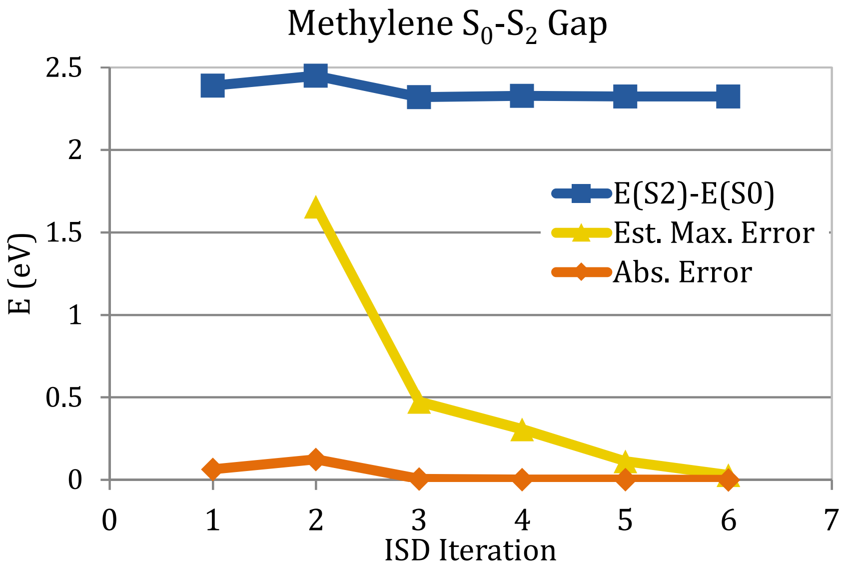


Figure 4.6 Convergence of singlet state gaps in CH_2 as a function of ISD iteration. a) S_0-S_1 gap. b) S_0-S_2 gap.

4.7 Conclusions

The proposed iterative submatrix diagonalization method is a promising new tool for diagonalizing the large Hamiltonian matrices of electronic structure theory. ISD converges to a desired energy threshold with control over eigenvalue errors and without using any additional parameters. The method is capable of handling ground and excited states, and can be used whenever natural orbitals are available. Therefore, ISD may be generally useful in other highly correlated electronic structure simulations, and likely can be extended for use with anharmonic vibrational structure methods.[57,58,59]

ISD has a reduced computational cost compared to the standard Davidson algorithm when converged to moderate thresholds (i.e. ~1 mHa for better than chemical accuracy) for ground states, and greatly reduced costs for gaps between electronic states. While additional studies are clearly needed to test the extensibility of ISD, the results presented herein and its performance in iFCI[12,13,14] are highly promising for application to the CI eigenvalue problem. The recent renewed interest in finding eigenvalues of large matrices, such as in FCI[10,11,29,32-35] suggests that new strategies such as ISD may prove useful. The next Chapter further explores strategies for treating large CI spaces with the heat-bath CI method, which focuses on selecting only those determinants which contribute significantly to the wave functions of interest.

4.8 References

1. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **17,** 87–94 (1975).
2. Liu, B. *Technical Report LBL-8158*. (1957).
3. Leininger, M. L., Sherrill, C. D., Allen, W. D. & Schaefer, H. F. Systematic Study of Selected Diagonalization Methods for Configuration Interaction Matrices. *J. Comput. Chem.* **22,** 1574–1589 (2001).
4. Roos, B. A new method for large-scale Cl calculations. *Chem. Phys. Lett.* **15,** 153–159 (1972).
5. Furche, F., Krull, B. T., Nguyen, B. D. & Kwon, J. Accelerating molecular property calculations with nonorthonormal Krylov space methods. *J. Chem. Phys.* **144,** 174105 (2016).
6. Parrish, R. M., Hohenstein, E. G. & Martínez, T. J. 'Balancing' the Block Davidson–Liu Algorithm. *J. Chem. Theory Comput.* **12,** 3003–3007 (2016).
7. Knowles, P. J. & Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **111,** 315–321 (1984).
8. Knowles, P. J. & Handy, N. C. A determinant based full configuration interaction program. *Comput. Phys. Commun.* **54,** 75–83 (1989).
9. Olsen, J., Roos, B. O., Jørgensen, P. & Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* **89,** 2185 (1988).
10. Olsen, J., Jørgensen, P. & Simons, J. Passing the one-billion limit in full configuration-interaction (FCI) calculations. *Chem. Phys. Lett.* **169,** 463–472 (1990).
11. Fales, B. S. & Levine, B. G. Nanoscale Multireference Quantum Chemistry: Full Configuration Interaction on Graphical Processing Units. *J. Chem. Theory Comput.* **11,**

4708–4716 (2015).
12. Zimmerman, P. M. Incremental full configuration interaction. *J. Chem. Phys.* **146,** 104102 (2017).
13. Zimmerman, P. M. Singlet–Triplet Gaps through Incremental Full Configuration Interaction. *J. Phys. Chem. A* **121,** 4712–4720 (2017).
14. Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *J. Chem. Phys.* **146,** 224104 (2017).
15. Sherrill, C. D. & Schaefer, H. F. The Configuration Interaction Method: Advances in Highly Correlated Approaches. *Adv. Quantum Chem.* **34,** 143–269 (1999).
16. Löwdin, P.-O. Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction. *Phys. Rev.* **97,** 1474–1489 (1955).
17. McWeeny, R. Some Recent Advances in Density Matrix Theory. *Rev. Mod. Phys.* **32,** 335–369 (1960).
18. Davidosn, E. R. Properties and Uses of Natural Orbitals. *Rev. Mod. Phys.* **44,** 451–464 (1972).
19. Barr, T. L. & Davidson, E. R. Nature of the Configuration-Interaction Method in Ab Initio Calculations. I. Ne Ground State. *Phys. Rev. A* **1,** 644–658 (1970).
20. Grev, R. S. & Schaefer, H. F. Natural orbitals from single and double excitation configuration interaction wave functions: their use in second-order configuration interaction and wave functions incorporating limited triple and quadruple excitations. *J. Chem. Phys.* **96,** 6850–6856 (1992).
21. Aquilante, F., Todorova, T. K., Gagliardi, L., Pedersen, T. B. & Roos, B. O. Systematic truncation of the virtual space in multiconfigurational perturbation theory. *J. Chem. Phys.* **131,** 34113 (2009).
22. Taube, A. G. & Bartlett, R. J. Frozen natural orbital coupled-cluster theory: Forces and application to decomposition of nitroethane. *J. Chem. Phys.* **128,** 164101 (2008).
23. Ivanic, J. & Ruedenberg, K. Deadwood in configuration spaces. II. Singles + doubles and singles + doubles + triples + quadruples spaces. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **107,** 220–228 (2002).
24. Neese, F. A spectroscopy oriented configuration interaction procedure. *J. Chem. Phys.* **119,** 9428–9443 (2003).
25. Landau, A., Khistyaev, K., Dolgikh, S. & Krylov, A. I. Frozen natural orbitals for ionized states within equation-of-motion coupled-cluster formalism. *J. Chem. Phys.* **132,** 14109 (2010).
26. Lu, Z. & Matsika, S. High-Multiplicity Natural Orbitals in Multireference Configuration Interaction for Excited States. *J. Chem. Theory Comput.* **8,** 509–517 (2012).
27. Lu, Z. & Matsika, S. High-Multiplicity Natural Orbitals in Multireference Configuration Interaction for Excited State Potential Energy Surfaces. *J. Phys. Chem. A* **117,** 7421–7430 (2013).
28. Epstein, P. S. The Stark Effect from the Point of View of Schroedinger's Quantum Theory. *Phys. Rev.* **28,** 695–710 (1926).
29. Nesbet, R. K. Configuration Interaction in Orbital Theories. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **230,** 312–321 (1955).
30. Møller, C. & Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **46,** 618–622 (1934).

31. Mayhall, N. J., Goldey, M. & Head-Gordon, M. A quasidegenerate 2nd-order perturbation theory approximation to RAS-nSF for excited states and strong correlations. *J. Chem. Theory Comput.* **10,** 589–599 (2014).
32. Cleland, D., Booth, G. H., Overy, C. & Alavi, A. Taming the First-Row Diatomics: A Full Configuration Interaction Quantum Monte Carlo Study. *J. Chem. Theory Comput.* **8,** 4138–4152 (2012).
33. Petruzielo, F. R., Holmes, A. A., Changlani, H. J., Nightingale, M. P. & Umrigar, C. J. Semistochastic Projector Monte Carlo Method. *Phys. Rev. Lett.* **109,** 230201 (2012).
34. Daday, C., Smart, S., Booth, G. H., Alavi, A. & Filippi, C. Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question. *J. Chem. Theory Comput.* **8,** 4441–4451 (2012).
35. Bytautas, L. & Ruedenberg, K. The Range of Electron Correlation between Localized Molecular Orbitals. A Full Configuration Interaction Analysis for the NCCN Molecule †. *J. Phys. Chem. A* **114,** 8601–8612 (2010).
36. Zimmerman, P. M., Bell, F., Goldey, M., Bell, A. T. & Head-Gordon, M. Restricted active space spin-flip configuration interaction: Theory and examples for multiple spin flips with odd numbers of electrons. *J. Chem. Phys.* **137,** 164110 (2012).
37. Huron, B., Malrieu, J. P. & Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **58,** 5745–5759 (1973).
38. Jeung, G.-H., Daudey, J.-P. & Malrieu, J.-P. Theoretical study of the lowest potential curves of NaK. *Chem. Phys. Lett.* **94,** 300–305 (1983).
39. Cimiraglia, R. & Persico, M. Recent advances in multireference second order perturbation CI: The CIPSI method revisited. *J. Comput. Chem.* **8,** 39–47 (1987).
40. Cimiraglia, R., Persico, M. & Tomasi, J. Ab initio study of the photodissociation of nitrosoalkanes and nitrosamines. *J. Am. Chem. Soc.* **107,** 1617–1622 (1985).
41. Schriber, J. B. & Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **144,** 161106 (2016).
42. Zhang, T. & Evangelista, F. A. A Deterministic Projector Configuration Interaction Approach for the Ground State of Quantum Many-Body Systems. *J. Chem. Theory Comput.* **12,** 4326–4337 (2016).
43. Tubman, N. M., Lee, J., Takeshita, T. Y., Head-Gordon, M. & Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **145,** 44112 (2016).
44. Nakatsuji, H. & Hirao, K. Cluster expansion of the wavefunction. Symmetry-adapted-cluster expansion, its variational determination, and extension of open-shell orbital theory. *J. Chem. Phys.* **68,** 2053–2065 (1978).
45. Holmes, A. A., Tubman, N. M. & Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **12,** 3674–3680 (2016).
46. Sharma, S., Holmes, A. A., Jeanmairet, G., Alavi, A. & Umrigar, C. J. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **13,** 1595–1604 (2017).
47. Andersson, K., Malmqvist, P. A., Roos, B. O., Sadlej, A. J. & Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **94,** 5483–5488

(1990).

48. Andersson, K., Malmqvist, P. & Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **96,** 1218–1226 (1992).
49. Pulay, P. A perspective on the CASPT2 method. *Int. J. Quantum Chem.* **111,** 3273–3279 (2011).
50. Ivanic, J. & Ruedenberg, K. Identification of deadwood in configuration spaces through general direct configuration interaction. *Theor. Chem. Acc.* **106,** 339–351 (2001).
51. Shao, Y. *et al.* Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **113,** 184–215 (2015).
52. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).
53. Feyereisen, M., Fitzgerald, G. & Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **208,** 359–363 (1993).
54. Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2,** 73–78 (2012).
55. Sherrill, C. D., Leininger, M. L., Van Huis, T. J. & Schaefer, H. F. Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (TZ2P) basis. *J. Chem. Phys.* **108,** 1040 (1998).
56. Zimmerman, P. M., Toulouse, J., Zhang, Z., Musgrave, C. B. & Umrigar, C. J. Excited states of methylene from quantum Monte Carlo. *J. Chem. Phys.* **131,** 124103 (2009).
57. Lin, C. Y., Gilbert, A. T. B. & Gill, P. M. W. Calculating molecular vibrational spectra beyond the harmonic approximation. *Theor. Chem. Acc.* **120,** 23–35 (2008).
58. Zimmerman, P. M. & Smereka, P. Optimizing Vibrational Coordinates To Modulate Intermode Coupling. *J. Chem. Theory Comput.* **12,** 1883–1891 (2016).
59. Molina, A., Smereka, P. & Zimmerman, P. M. Exploring the relationship between vibrational mode locality and coupling using constrained optimization. *J. Chem. Phys.* **144,** 124111 (2016).

Chapter 5: Excited States of Methylene, Polyenes, and Ozone from Heat-Bath Configuration Interaction

5.1 Abstract

The electronically excited states of methylene, ethylene, butadiene, hexatriene, and ozone have long proven challenging due to their complex mixtures of strong and weak correlations. To provide accurate benchmark energies for these states, semistochastic heat-bath configuration interaction is herein used to approximate full configuration interaction (FCI) energies. A new metric, the fraction of correlation energy obtained by the perturbative correction (f_{PT}), is shown to be a good indicator of convergence. The presented butadiene/ANO-L-pVTZ and hexatriene/ANO-L-pVDZ computations are the largest FCI-level simulations to date, with FCI dimensions exceeding 10^{35} and 10^{38} determinants. The best heat-bath estimates of the vertical excitation energies in butadiene/ANO-L-pVTZ are 6.73 eV for 2^1A_g, and 6.37 eV for 1^1B_u (6.74 and 6.57 eV with ANO-L-pVDZ). The same gaps in hexatriene/ANO-L-pVDZ are estimated to be 5.73 and 5.61 eV, respectively.

5.2 Introduction

Full configuration interaction (FCI) provides the theoretically exact electronic energy in a given one-particle basis set, making FCI calculations the benchmarks against which other quantum chemistry methods are evaluated.[1–7] The exponential increase in Hamiltonian dimension with increasing system size, however, means that traditional FCI benchmarks are not easily achievable for molecular systems larger than diatomics The ISD algorithm presented in Chapter 4 managed to reduce the CI spaces and time necessary to obtain FCI-quality wave functions by about one order of magnitude. However, to dramatically increase the size of molecules that can be treated at the FCI level, more drastic reductions in CI space will be required. Alternatives to ISD have seen impressive progress in recent years, producing FCI-quality energies for much larger systems with such algorithms as incremental FCI (iFCI),[8–10] density matrix renormalization group (DMRG),[11–15] and FCI quantum Monte Carlo (FCIQMC).[16,17]

These advances have also led to a revival of interest in the concept of select CI (SCI). SCI iteratively builds a subspace Hamiltonian[18] composed of the most important determinants of the wave function, which are generally selected based on perturbative coefficients. Usually, perturbative energy corrections to the lowest energy eigenvectors are used to correct for deficiencies in the variational subspace, leading to SCI+PT levels of theory. SCI+PT methods were first popularized by the CIPSI method, configuration interaction by perturbatively selecting iteratively.[19–24] CIPSI becomes prohibitively expensive once the variational space grows large enough, as the algorithm generates and evaluates all determinants connected to this space. The first major improvement over CIPSI was adaptive sampling CI (ASCI),[25] a deterministic variant of FCIQMC that truncates the sum in the perturbative coefficient expression, reducing the cost of evaluating a determinant's importance. Despite this improvement, the requirement to generate and evaluate all connected determinants remained a bottleneck. Heat-bath CI (HCI)[26] was designed to eliminate this problem by only considering determinants which pass its new, more simple importance measure. HCI was further improved by stochastic evaluation of the PT energy in semistochastic HCI (SHCI),[27] which eliminated the need to store a long list of perturbative determinants in memory.

Because FCI is a benchmark-level method with extreme cost, results from FCI are typically most useful in evaluating other electronic structure methods. Unfortunately, benchmarks for polyatomics have largely been unavailable, and thus the subtle effects of correlations amongst numerous electrons have not been accurately quantified. With SHCI, however, systems of substantial size—at least compared to prior algorithms—can now be treated at the FCI level. We utilize the code of Umrigar et al.[27] to push SHCI to its limits, both by maturing the code through rigorous testing as well as applying it to large systems. Specifically, our aim is to provide highly accurate benchmarks for electronically excited states of the challenging polyatomics in Figure 5.1. In this work, methylene is presented as the first test case due to its small size yet challenging electronic structure.[28] Ozone is also investigated as a chemically relevant case due to its role in upper atmosphere chemistry.[29]

Finally, ethylene, butadiene, and hexatriene are examined to provide benchmarks for the first few members of the polyene series, long studied for their role as prototypical organic conducting polymers. The low-lying valence excited states, 2^1A_g and 1^1B_u,[30–38] are especially challenging to characterize and have long been a target of accurate electronic structure theories. With 10^{35} and 10^{38} determinants in the FCI space of butadiene/ANO-L-pVTZ and hexatriene/ANO-L-pVDZ, respectively, FCI-level computations are incredibly challenging, and will push SHCI towards its limit.

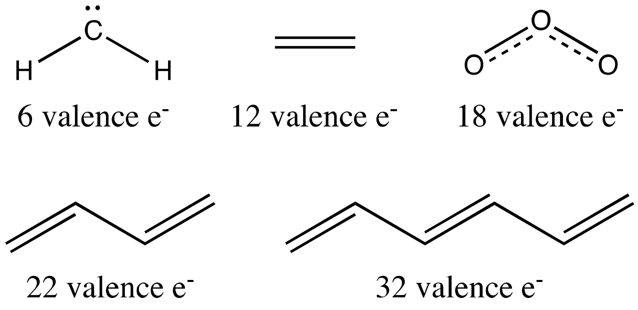


Figure 5.1 Molecules investigated with SHCI.

This article is organized as follows. In Sections 5.3 and 5.4, the SHCI methodology for excited state and error estimation strategies are laid out. Computational Details are laid out in Section 5.5 and Section 5.6 presents SHCI results for methylene, ozone, ethylene, butadiene, and hexatriene. Small systems in which benchmark values can be obtained (methylene, ethylene, and ozone/cc-pVDZ) are first examined to obtain insights into the convergence behavior of SHCI. These observations are then used to estimate the convergence level of ozone/cc-pVTZ, butadiene, and hexatriene results. Section 5.7 provides conclusions and an outlook on the SHCI method for excitation energy benchmarking.

5.3 Semistochastic Heat-Bath Configuration Interaction

As the HCI and SHCI algorithms have been described in detail,[26,27] only a brief overview will be given here. The HCI algorithm can be divided into variational and perturbational stages, each of which selects determinants through threshold values, ϵ_{var} and ϵ_{PT}, respectively. The variational space (V) contains the determinants included in the current subspace, and the connected space (C) contains all determinants connected by single or double excitations to V.

The variational stage iteratively adds determinants to V by

1. Adding all determinants connected to the current V that pass the importance criterion $|H_{ai}\max(|c_i|)| > \epsilon_{var}$, where $\max(|c_i|)$ is the largest coefficient currently associated with determinant *i* for the states of interest.
2. Constructing the Hamiltonian and solving for the roots of interest, in the basis of all determinants in the newly expanded V.
3. Repeating 1-2 until convergence.

The HCI importance measure is much easier to evaluate than the full first-order perturbative estimate of a determinant's wave function coefficient, $c_a^{(1)} = \left|\frac{\sum_i H_{ai}c_i}{E-H_{aa}}\right|$, used in other SCI+PT methods such as CIPSI[19] and ASCI.[25] This simplification is possible because $c_a^{(1)}$ is dominated by a few terms due to large variations in H_{ai} and c_i, which span multiple orders of magnitude, whereas $E - H_{aa}$ varies by less than 1 order of magnitude.[26] Using the HCI importance measure, doubles excitations can be sorted by the magnitude of their Hamiltonian coupling elements, and excitations with small H_{ai} values can be bypassed entirely.

After convergence of V, signified by few additional determinants or small variational energy changes, a second-order perturbative energy correction for the variational energy E is calculated by

$$\Delta E^{(2)} = \sum_k \frac{(\sum_i H_{ki}c_i)^2}{E - H_{kk}}$$

where *k* runs over all determinants in C, and *i* ranges over all determinants in V. Similar to the variational stage, the perturbation only considers the determinants connected to the final V space that have an importance measure greater than a parameter ϵ_{PT}, which is typically much smaller than ϵ_{var}. The storage of the entire space of determinants used in $\Delta E^{(2)}$ becomes a memory bottleneck for larger systems. This memory bottleneck can be sidestepped by calculating the second-order perturbation correction semistochastically, as in the SHCI algorithm.[27] In SHCI, a deterministic energy correction is first calculated with a $\epsilon_{PT}^{d} > \epsilon_{PT}$, and the error from using the larger ϵ_{PT}^{d} is then calculated stochastically by taking the difference of the second-order corrections done with ϵ_{PT}^{d} and ϵ_{PT}.[27] Samples are taken until the statistical error falls below a specified threshold.

5.4 Error Analysis

The target accuracy for absolute or relative energies is typically chosen to be either within 1 mHa or 1.6 mHa (1 kcal/mol, representing chemical accuracy) of the FCI limit. A major obstacle in examining systems as large as hexatriene (FCI dimension 10^{38}) is accurately determining the extent of convergence relative to the FCI limit. Straightforward strategies, such as examining changes in energy between increasingly accurate runs or extrapolating energies to the FCI limit, can be unreliable unless one is extremely close to the FCI limit. For instance, varying ϵ by a small amount may produce a small change in total energy, but this does not necessarily indicate convergence to the true FCI energy. Furthermore, extrapolation procedures can result in significantly different energies depending on the available data and the chosen fitting equation. As neither of these measures is a certain signal for convergence in the larger systems for which we cannot closely approach the FCI limit (ozone/cc-pVTZ, butadiene, and hexatriene), this work utilizes a new convergence metric, the fraction of correlation energy obtained by the perturbative correction (f_{PT}), to support the observed convergence of SHCI results.

f_{PT} can be used as a convergence metric because the error in total energy arises only from the perturbative correction, since the variational treatment is exact. Therefore, if the perturbative correction is less than 1 mHa, the variational energy itself is converged to within 1 mHa (we assume small ϵ_{PT} such that the perturbative correction itself is converged). Such small perturbative corrections are often unobtainable, however, so an alternative accuracy measure is required for general application. For this alternative, we use f_{PT}, which utilizes the distribution of the recovered correlation energy between variational and perturbative contributions, to indicate closeness to the FCI limit. As f_{PT} falls, so too should the error, and there may be a fraction at which total or relative energies are generally converged to the FCI limit. This claim is examined and shown to be empirically true for relative energies in the smaller systems for which benchmark values can be obtained.

To use f_{PT} for excited states, which are not well described by single electronic configurations, a slightly different definition of correlation energy is required. Instead of taking the correlation energy as the energy difference between Hartree-Fock and FCI, we utilize a minimal complete active space (CAS) wave function in place of Hartree-Fock, resulting in $E_{corr} = E_{FCI} - E_{CAS}$. This leads to a well-defined f_{PT} for all states, with f_{PT} values for relative energies defined as the average f_{PT} of the two involved states.

5.5 Computational Details

SHCI is implemented in Fortran90, parallelized using MPI, and makes use of symmetry.[27] Convergence of the variational space is signaled when the number of new determinants added is less than 0.001%, of the current variational space, or when the change in energy is less than $1 \cdot 10^{-5}$ Ha. The ANO-L-pVXZ basis sets[39] are used for ethylene, butadiene, and hexatriene. For methylene, the aug-cc-pVQZ[40,41] basis is used. For ozone, cc-pVDZ and cc-pVTZ[40] basis sets are used to examine f_{PT} and facilitate direct comparisons with benchmark values.[29] ϵ_{var} is varied for each system to show convergence of energies or energy gaps with the series ϵ_{var} ={$7 \cdot 10^{-X}$, $5 \cdot 10^{-X}$, $3 \cdot 10^{-X}$, $1 \cdot 10^{-X}$}, where X={2, 3, 4, 5, 6}. The range varies by system due to computational resources and convergence behavior, and these details are presented below. For all calculations, ϵ_{PT} is set to 10^{-7} Ha, which provides converged perturbation corrections.[26,27] The statistical error of the stochastic perturbative correction is generally set to 0.05 mHa, with the larger hexatriene/ANO-L-pVDZ and butadiene/ANO-L-pVTZ calculations set to 0.1 mHa. All calculations use the frozen-core approximation and state-specific natural orbitals derived from variational HCI calculations for each system at tight ϵ_{va} . In butadiene, hexatriene, and ozone, the HCI-generated natural orbitals for the two $A_{g/1}$ symmetry states are state-averaged with equal weighting. Methylene's geometries are FCI/TZVP quality, taken from Sherrill et al.[28] Ethylene, butadiene, and hexatriene geometries are optimized at the RI-MP2/cc-pVQZ level. Ozone's geometries are CASSCF(18,12)/cc-pVQZ quality, taken from Theis et al.[29] For comparisons to coupled cluster theories, the same geometries and basis sets are used with Q-Chem 4.0's[42] implementation[43] of CR-EOM-CC(2,3)D.[44] Active spaces for minimal CAS-CI computations (used in f_{PT}) were chosen based on the coefficients of the most accurate calculations for each system. Methylene used a CAS(2,4), ozone a CAS(6,4), and the polyenes the full π-systems.

5.6 Results

SHCI was used to investigate the low-lying valence states of methylene, ethylene, ozone, butadiene, and hexatriene. These systems contain 6, 12, 18, 22, and 32 valence electrons, all of which will be correlated with the full virtual orbital space. With no restriction on the CI excitation level, the massive combinatorial spaces of CI determinants make it difficult to guarantee convergence in systems with over 20 valence electrons. Therefore, the variational energies (E_{var}), perturbative corrections (E_{PT}), and total (E_{SHCI}) energies from a progression of ϵ_{var} calculations down to benchmark levels will first be examined to determine meaningful convergence metrics for the larger molecules.

5.6.1 Estimating Chemical Accuracy from f_{PT}

SHCI can obtain benchmark energies fully converged to the FCI limit for the smaller test systems: methylene with aug-cc-pVQZ basis, ethylene with ANO-L-pVDZ and ANO-L–pVTZ bases, and ozone with cc-pVDZ. The accuracy of these benchmarks is confirmed via either sub-mHa perturbative corrections, or consistent extrapolations to the FCI limit. For methylene with aug-cc-pVQZ and ethylene with ANO-L-pVDZ, perturbative corrections are sub-mHa. For ethylene with ANO-L-pVTZ and ozone with cc-pVDZ, perturbative corrections are 1.4 – 2.6 mHa, allowing accurate extrapolations to FCI. The small E_{PT} indicates closeness to the FCI limit, which minimizes errors in extrapolation, allowing consistent energies to be returned with various fitting equations and data points. This is indeed the case (Appendix D.2 and D.3), with all extrapolated energies being within 0.4 mHa of the most accurate obtained E_{SHCI}.

Methylene's convergence progression is well behaved due to having only six valence electrons (Appendix D.1). For ethylene, the relationship between f_{PT} and absolute error is presented in Figure 5.2 for ANO-L-pVTZ, while ANO-L-pVDZ results show similar trends, as seen in the Appendex D.2 and Table 5.1. Ethylene's absolute energies require an average f_{PT} of 14.0% and 12.1% to become within 1.6 and 1 mHa of the converged SHCI limit, respectively, whereas relative energies reach errors of 1.6 and 1 mHa with average f_{PT} values of 14.8% and 10.9%. The relative energies require slightly lower f_{PT} (i.e. tighter ϵ_{var}) to converge to sub-mHa accuracies. This is somewhat surprising as one expects energy gaps to converge faster than absolute energies due to cancellation of errors. Closer examination shows that this behavior is due to the non-variational character of the perturbative corrections, which allow E_{SHCI} to converge to the FCI limit from below or above. This situation occurs for ethylene's 1^3B_{1u} and 1^1A_g states, which converge to their FCI limits from above and below respectively, leading to slower convergence of the relative energies.

Table 5.1 Values of f_{PT} where errors for absolute and relative energies have fallen below 1.6 or 1 mHa.

Required f_{PT} for Convergence (%)					
Energy Type		Absolute		Relative	
Error Threshold (mHa)		1.6	1	1.6	1
Methylene/aug-cc-pVQZ		10.0	6.6	12.4	8.2
Ethylene	ANO-L-pVDZ	12.9	10.3	17.1	10.4
	ANO-L-pVTZ	14.0	12.1	13.8	10.9
Ozone/cc-pVDZ		4.4	2.8	21.3	13.8
Average		10.3	8.0	16.4	10.8

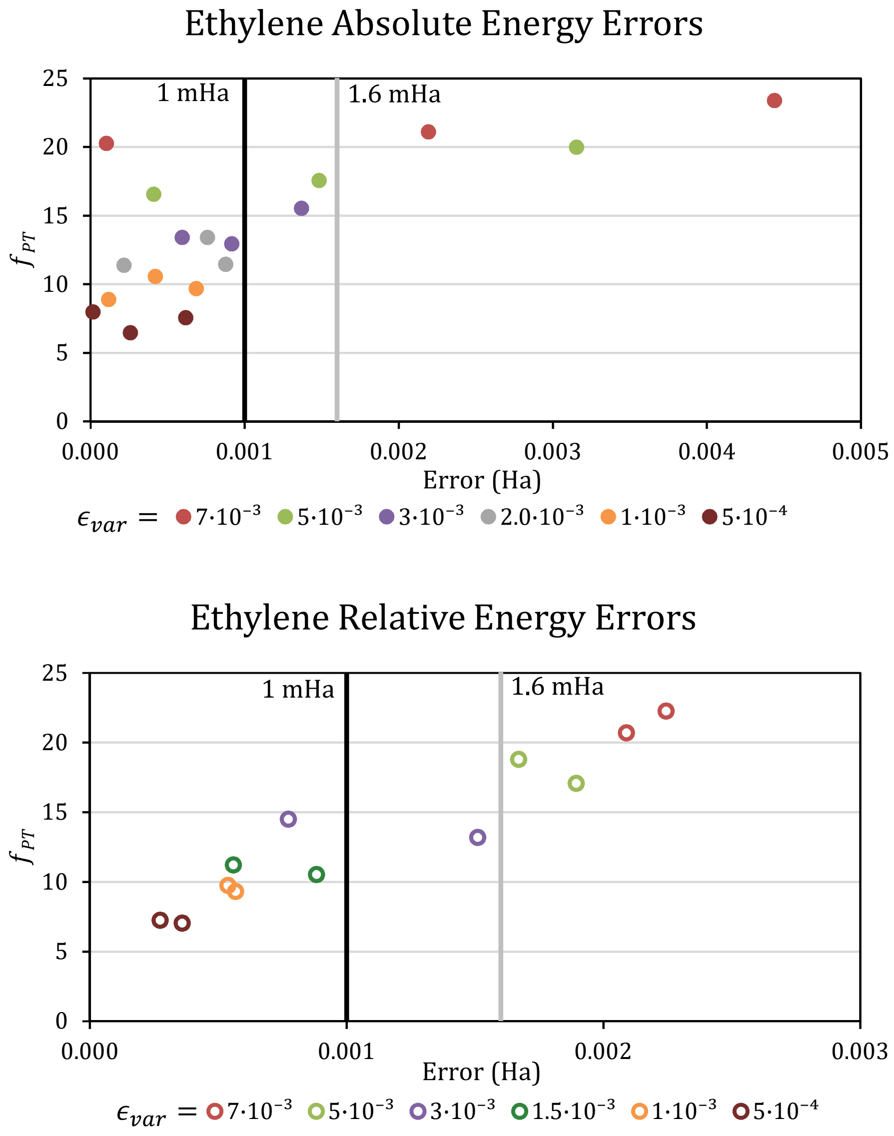


Figure 5.2 The relationship between f_{PT} and absolute error for relative and absolute energies in ethylene/ANO-L-pVTZ.

Ozone in the cc-pVDZ basis and at its ground-state geometry is the next benchmark system. The most accurate run was accomplished with $\epsilon_{var} = 7 \cdot 10^{-6}$ Ha, giving a 2^1A_1-1^1A_1 gap of 4.10 eV. With SHCI, the gap is converged to within 1 mHa when treating only 10^4 determinants out of the 10^{16} in the FCI space. Figure 5.3 graphs the relationship between f_{PT} and errors as in ethylene, where it can be seen that a small f_{PT} of 2.8% is necessary to converge absolute energies to sub-mHa errors. The relative energies, however, converge more quickly, with f_{PT} values of 21.3% and 13.8% needed to obtain errors under 1.6 and 1 mHa, respectively.

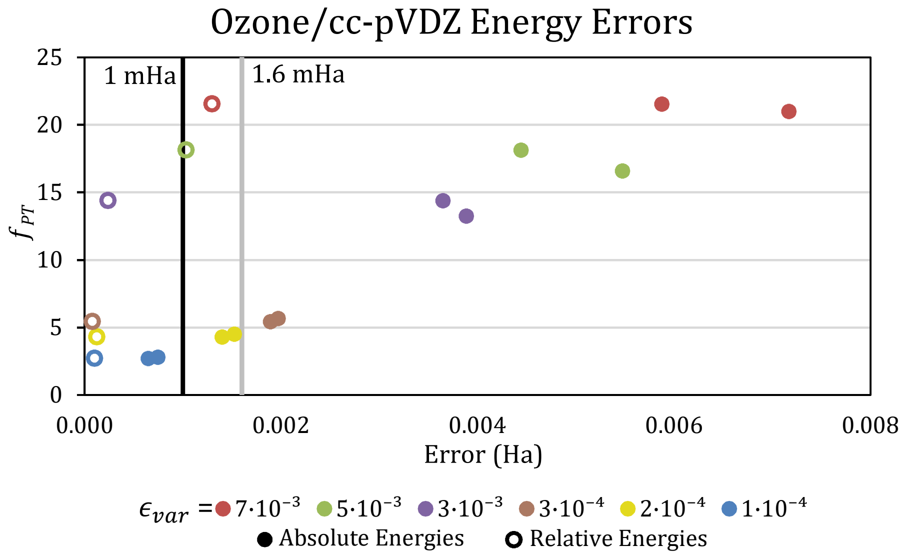


Figure 5.3 The relationship between f_{PT} and absolute error for both relative and absolute energies in ozone/cc-pVDZ at the equilibrium geometry.

The values of f_{PT} that indicate convergence of E_{SHCI} are compiled in Table 5.1. Absolute energy errors are inconsistently estimated by f_{PT}, as seen by the large difference in f_{PT} values for ozone compared to methylene and ethylene. The relationship between f_{PT} and relative energies is better behaved, with only minor variations in the f_{PT} quantities signaling tight convergence. Relative energies are consistently converged to within 1 mHa with an average f_{PT} = 10.8% and to within 1.6 mHa with an average f_{PT} = 16.4%. The required f_{PT} for convergence appears to rise as the number of correlated electrons increases. Thus, it is possible that in the larger molecules, relative energies will converge at slightly higher f_{PT} than needed for the benchmark systems. Regardless, this work will utilize conservative values of f_{PT} to signal convergence, with f_{PT} values below 10% and 15% taken to signal errors in relative energies below 1 and 1.6 mHa respectively.

5.6.2 Methylene

Methylene is a prototypical test case for advanced electronic structure methods, being small enough to be amenable to canonical FCI benchmarks, yet still requiring accurate treatment of dynamic and static correlations for correct excitation energies.[45–51] The four lowest lying states of methylene vary in spin and spatial symmetry, 1^3B_1, 1^1A_1,

1^1B_1, and 2^1A_1. With only six valence electrons, SHCI can be used with the large aug-cc-pVQZ basis. With $\epsilon_{var} = 1 \cdot 10^{-5}$ Ha, perturbative corrections were less than 0.01 mHa, indicating strict convergence of the SHCI values to the FCI limit. The Appendix D.1 shows methylene's adiabatic energy gaps are converged to sub-mHa accuracy at $\epsilon_{var} = 3 \cdot 10^{-3}$ Ha, so $\epsilon_{var} = 1 \cdot 10^{-5}$ Ha represents tighter convergence than needed to reach 1 mHa accuracy. Table 5.2 shows the most accurate SHCI adiabatic energy gaps differ from experiment by about 0.01 eV.[28] Comparing canonical FCI in the TZ2P basis with SHCI in the larger aug-cc-pVQZ basis shows differences of up to 0.144 eV, suggesting that large basis sets are necessary to fully describe correlation in methylene. This was first hinted at by diffusion Monte Carlo (DMC) results,[48] which are less sensitive to basis set, that agree with SHCI to within about 0.02 eV. CR-EOMCC(2,3)D relative energies are also converged to within 1 mHa of the benchmark SHCI values, indicating that high-level coupled cluster calculations can correlate six electrons sufficiently to obtain FCI-quality energy gaps.

Table 5.2 Methylene total and relative energies.

State	SHCI[a] (Ha)				
1^3B_1	-39.08849(1)				
1^1A_1	-39.07404(1)				
1^1B_1	-39.03711(1)				
2^1A_1	-38.99603(1)				
Gaps	SHCI[a] (eV)	CR-EOMCC(2,3)D (eV)	FCI (eV)[b]	DMC (eV)[c]	Exp (eV)
1^1A_1-1^3B_1	0.393	0.415	0.482	0.406	0.400[d]
1^1B_1-1^3B_1	1.398	1.422	1.542	1.416	1.411[e]
2^1A_1-1^3B_1	2.516	2.499	2.674	2.524	-

[a] E_{SHCI} at $\epsilon_{var} = 1 \cdot 10^{-5}$
[c] FCI/TZ2P results from reference [28]
[c] Diffusion Monte Carlo results from reference [48]
[d] References [28], [52]
[e] References [28], [53]

5.6.3 Ethylene

Ethylene is a prototypical benchmark system for electronic excitations, including a difficult-to-characterize 1^1B_u state. Although the 1^1B_u state is qualitatively well described by a π-π* excitation, a quantitative description requires a thorough accounting of dynamic correlation between π and σ electrons.[54–56] Here, SHCI is applied to the low-lying valence states, 1^1A_g, 1^1B_{1u}

and 1^3B_{1u}, in the ANO-L-pVTZ basis. Table 5.3 shows that SHCI total and relative energies compare favorably with previous FCIQMC[57] and iFCI[9] results (Appendix D.2 gives results in the smaller ANO-L-pVDZ basis). Table 5.3 also indicates that coupled cluster methods must account for more than triples excitations in order to accurately correlate this system, as CR-EOMCC(2,3)D relative energies show errors greater than 1.6 mHa with respect to the SHCI benchmark values. With SHCI, absolute energies within 1 mHa of the FCI limit were obtained at $\epsilon_{var} = 1 \cdot 10^{-3}$ Ha, where only 10^5 variational determinants were considered out of a FCI space of 10^{18}. These results suggest that polyatomics with up to 12 valence electrons and triple-zeta basis sets can now be considered unchallenging for FCI-level approximation. SHCI benchmarks agree with prior simulations which indicate the 1^1B_{1u} vertical excitation does not correspond to experimental band maxima.[58]

Table 5.3 SHCI ethylene/ANO-L-pVTZ energies compared to literature values.

State	SHCI[a] (Ha)		FCIQMC (Ha)[b]		
1^1A_g	-78.4381(1)		-78.4370(2)		
1^1B_{1u}	-78.1424(1)		-78.1407(3)		
1^3B_{1u}	-78.2693(1)		-		
Gap	SHCI[a] (eV)	CR-EOMCC(2,3)D (eV)	FCIQMC (eV)[b]	iFCI (eV)	Exp (eV)
1^1B_{1u}-1^1A_g	8.05	8.25	8.06[b]	-	7.66[d]
1^3B_{1u}-1^1A_g	4.59	4.76	-	4.64[c]	4.3-4.6[e]

[a] E_{SHCI} at $\epsilon_{var} = 1 \cdot 10^{-5}$

[b] FCIQMC/ANO-L-pVTZ results from reference 57

[c] iFCI/cc-pVTZ results from reference 9

[d] Experimental band maximum from reference 58

[e] Experimental band maxima from references 59–61

5.6.4 Ozone

Ozone's potential energy surfaces have held great interest due to its role in atmospheric chemistry.[62] An interesting feature of these surfaces predicted by computational studies is the existence of a metastable ring geometry on the ground state surface.[63] A lack of experimental evidence for such a species has fueled multiple studies of the pathway leading to the ring species over the years.[64–69] The most recent such study by Ruedenberg et al. utilizes multi-reference CI with up to quadruple excitations,[29] expending considerable effort on selecting and justifying an active space. To provide an accurate picture at critical points along the theorized pathway with even treatment of all valence electrons, SHCI is applied to ozone's 2^1A_1-1^1A_1 gap with cc-pVTZ at the three geometries of interest shown in Figure 5.4: the equilibrium geometry (termed the open ring minimum (OM)), the hypothetical ring minimum (RM), and the transition state (TS) between these two.

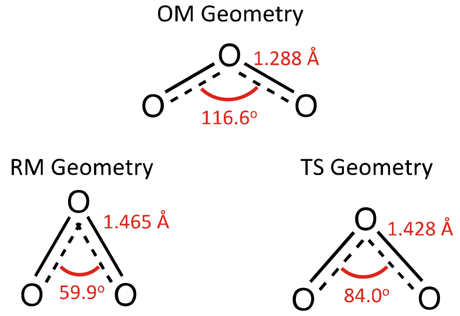


Figure 5.4 Ozone geometries. OM is the open minimum, ground state equilibrium geometry. RM is the ring minimum, a metastable geometry predicted by theory. TS is the transition state between OM and RM.

.

Ozone in the cc-pVTZ basis is too large to reach sub-mHa perturbative corrections, as computations with the necessary ϵ_{var} would exceed available memory. The E_{PT} corrections for the best available SHCI calculation, at $\epsilon_{var} = 3 \cdot 10^{-5}$ Ha, range from 15.8-25.8 mHa, and the extrapolated energies differ by up to 2.6 mHa from the best SHCI values (Appendix D.3), indicating that sub-mHa accuracy in total energies was not obtained. f_{PT} values of 2.8% and under, however, are recovered for all geometries at $\epsilon_{var} = 3 \cdot 10^{-5}$ Ha, which is well below the metric established for chemical accuracy of relative energies (15%). Furthermore, graphing the energy gaps at various ϵ_{var} (Figure 5.5) shows little variance as the tightest convergence limit is approached. The RM data set in Figure 5.5 is truncated above $\epsilon_{var} = 5 \cdot 10^{-4}$ (i.e. $-\log \epsilon_{var} = 3.3$) because the 2^1A_1 wave function is qualitatively different at looser ϵ_{var}, leading to large changes in the 2^1A_1-1^1A_1 energy gap. This serves as a cautionary case, demonstrating that qualitatively correct wave functions are not guaranteed at loose ϵ_{var}, as important groups of determinants may be left out. This issue was only seen with a relatively large number of electrons correlated in a moderately large basis. In these cases, we recommend $\epsilon_{var} \leq 1 \cdot 10^{-4}$ to ensure reasonable convergence.

In Table 5.4, the SHCI energy gaps are compared to Ruedenberg et al's MRCI results.[29] SHCI results mostly resemble MRCI estimates, except for the RM geometry, where the gaps differ by more than 1 eV. Due to the relatively tight convergence, these SHCI results should be taken as the current best 2^1A_1-1^1A_1 vertical excitation energies at the OM, RM, and TS geometries. SHCI therefore also allows some insight into the nature of the RM species. Along the 1^1A_1 potential surface, the RM and TS geometries lie 1.30 eV and 2.40 eV, respectively, above the OM geometry. This large barrier suggests that electronic excitations in ozone are likely required to reach RM, but the RM species should be relatively stable with a 1.10 eV barrier to return to the OM geometry. To obtain these results, 10^8 variational determinants out of 10^{23} in the FCI space were required, but no active space selection was needed.

Table 5.4 SHCI ozone/cc-pVTZ 2^1A_1-1^1A_1 gaps compared with Ruedenberg et al. All values in eV.

Geometry	SHCI[a]	MRCI (SDTQ)[b]
OM	4.13	3.54-4.63
RM	6.16	7.35-8.44
TS	0.04	0.05-0.16

[a] E_{SHCI} at $\epsilon_{var} = 3 \cdot 10^{-5}$ Ha

[b] From reference [29]

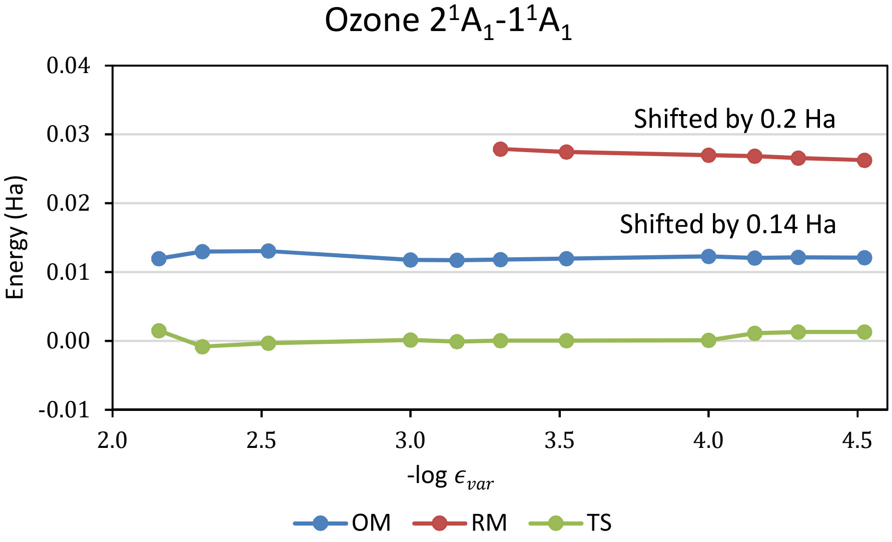


Figure 5.5 2^1A_1-1^1A_1 energy gaps versus -log ϵ_{var} for ozone/ANO-L-pVTZ at OM, RM, and TS geometries.

5.6.5 Shorter Polyenes: Butadiene and Hexatriene

Butadiene and hexatriene are part of the polyene series, long studied for their role as prototypical organic conducting polymers. The spacing of the low-lying valence excited states has proven especially challenging to electronic structure methods.[30–38] In butadiene and hexatriene, SHCI is thus applied to the 1^1A_g, 1^1B_{1u}, 1^3B_{1u}, and 2^1A_g states to provide accurate benchmarks. Butadiene and hexatriene are of special interest because their 1^1B_{1u} and 2^1A_g states are nearly degenerate, resulting in conflicting reports of state ordering at lower levels of theory. These systems are too large for the routine application of FCI-level methods, although limited FCIQMC[57] and DMRG[15] studies have been performed on butadiene.

5.6.5.1 Butadiene

Butadiene (C_4H_6) has FCI spaces of 10^{26} and 10^{35} determinants in the ANO-L-pVDZ and pVTZ basis sets, respectively, putting it at the edge of accessibility for modern FCI-quality approximations.[15,57] In the triple-zeta basis at $\epsilon_{var} = 3 \cdot 10^{-5}$ Ha, the resulting variational space of 10^8 determinants leads to the perturbative corrections of ~60 mHa and $f_{PT} < 10\%$ (Table 5.5). In butadiene, all states approach the FCI limit from above, allowing cancellation of errors to produce accurate energy gaps for f_{PT} above the recommended 10%. Figure 5.6 provides further evidence that the energy gaps are converged, as they vary by less than 1 mHa as the tightest convergence level ($\epsilon_{var} = 3 \cdot 10^{-5}$) is approached.

For the 1^1A_g ground state in an ANO-L-pVDZ basis, SHCI results can be compared in more detail to related FCI-level approximations in the literature. ANO-L-pVDZ's 1^1A_g state energy of -155.5528 is within 4 mHa of the most accurate estimates, a DMRG result of -155.5572[15] and an iFCI value of -155.5567 Ha.[10] Linear extrapolation of the energies slightly overshoots these targets, producing -155.5592 Ha (Appendix D.4). These results affirm that the absolute energy errors are not well predicted by the f_{PT} metric (c.f. Table 5.1), which is 6.3% at the tightest ϵ_{var} ($3 \cdot 10^{-5}$).

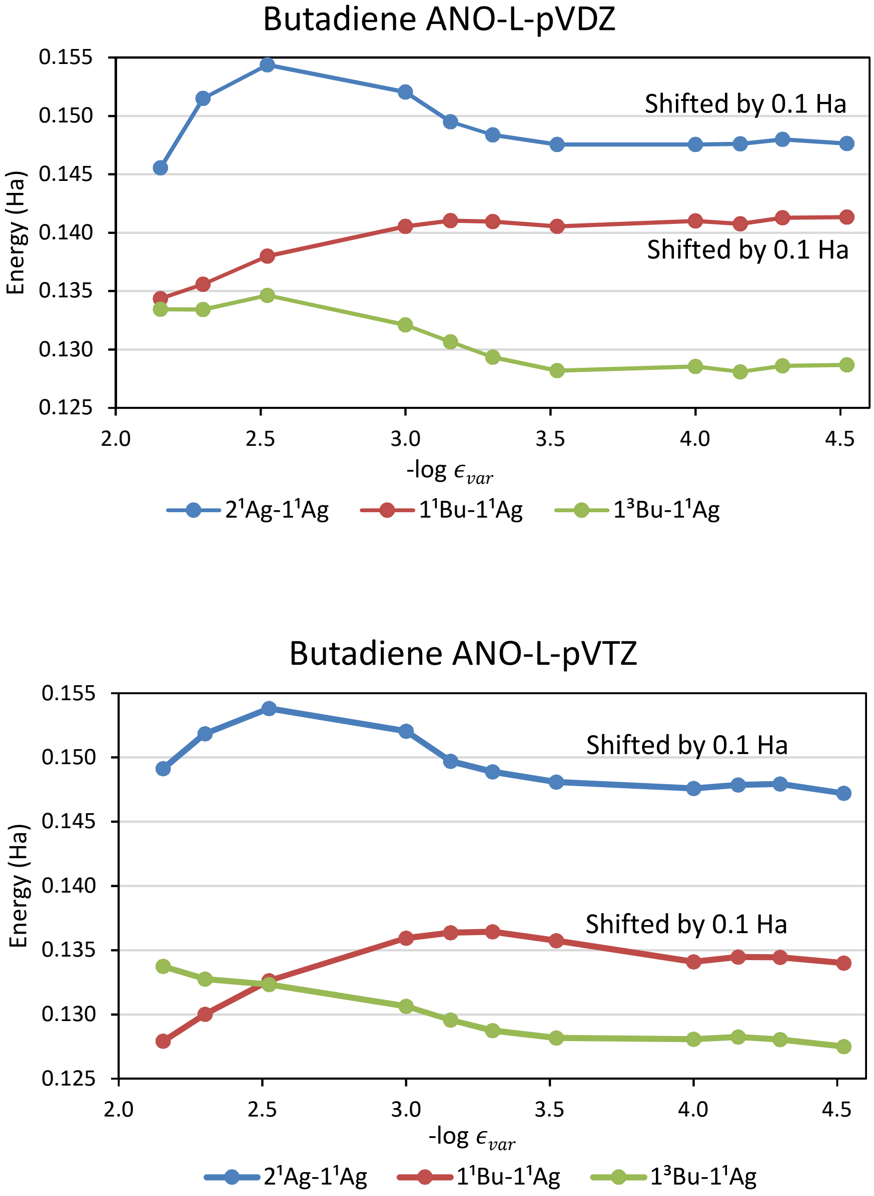


Figure 5.6 Energy gaps versus -log ϵ_{var} for butadiene with ANO-L-pVDZ and ANO-L-pVTZ basis sets. The energy gaps are well converged by $\epsilon_{var} = 1 \cdot 10^{-4}$ Ha.

Table 5.5 shows that SHCI butadiene energy gaps are similar to previous high-level theoretical calculations. For ANO-L-pVTZ, the calculated 1^3B_u-1^1A_g gap at is just 0.03 eV away from that computed by iFCI, and the 1^1B_u-1^1A_g gap is within 0.01 eV of FCIQMC. Since the 2^1A_g-1^1A_g energy gap does not currently have other FCI-level benchmarks, the SHCI values represent the current best theoretical estimate for the 2^1A_g-1^1A_g energy gap in butadiene, 6.73 eV. The small gap between the 2^1A_g and 1^1B_u states is consistent with recent theoretical[70] and experimental[71] investigations demonstrating ultrafast population transfer from 1^1B_u to 2^1A_g, which implies close proximity of the two states. SHCI's relative ordering of the 2^1A_g / 1^1B_u states agrees qualitatively with EOM-CC results,[72] but the 2^1A_g-1^1A_g gaps differ by 0.36 eV, indicating that the doubly-excited character of 2^1A_g was treated inadequately in the EOM-CC approach. As in ethylene, the 1^1B_u-1^1A_g gap differs from experiment by 0.45 eV, agreeing with prior indications that experimental band maxima of butadiene do not correspond to the vertical excitation energy.[72]

Table 5.5 Comparison of SHCI butadiene energy gaps with literature values. All values in eV.

Basis Set	Gap	f_{PT} (%)	Best SHCI[a]	EOM-CCSDTQ/MBE[b]	FCIQMC	iFCI	Exp
ANO-L-pVDZ	2^1A_g-1^1A_g	7.6	6.74	-	-	-	-
	1^1B_u-1^1A_g	6.6	6.57	-	6.53[c]	-	5.92[f]
	1^3B_u-1^1A_g	6.9	3.5	-	3.45[d]	3.45[e]	3.22[g]
ANO-L-pVTZ	2^1A_g-1^1A_g	8.1	6.73	6.39	-	-	-
	1^1B_u-1^1A_g	7.9	6.37	6.21	6.38[d]	-	5.92[f]
	1^3B_u-1^1A_g	7.8	3.47	-	-	3.44[e]	3.22[g]

[a] E_{SHCI} at $\epsilon_{var} = 3 \cdot 10^{-5}$ Ha for both ANO-L-pVDZ/TZ basis sets
[b] EOM-CC results from reference [72]
[c] FCIQMC/ANO-L-pVDZ result from reference [57]
[d] FCIQMC/truncated ANO-L-pVTZ result from reference [57]
[e] iFCI 6-31G* and cc-pVTZ results from reference [9]
[f] Experimental band maxima from references [73–75]
[g] Electron impact band maximum from reference [76]

5.6.5.2 Hexatriene

Hexatriene is at the current frontier of FCI-level computations, with a demanding FCI space of 10^{38} in a double-zeta basis. Only one other algorithm, iFCI,[9] has approached FCI energies for such a large polyatomic. iFCI has estimated the singlet-triplet gap for hexatriene, but is not immediately applicable to singlet excited states. The current highest level computational estimates of valence energy gaps in hexatriene are therefore presented here. Hexatriene's relative energies have f_{PT} values under 14.9%, obtained by treating 10^8 determinants variationally at $\epsilon_{var} = 5 \cdot 10^{-5}$ Ha. This value is close to the f_{PT} threshold of 15% for chemical accuracy, but since all energies all converge to the FCI limit from above, improved cancellation of errors is expected. Indeed, the hexatriene energy gaps shown in Figure 5.7 indicate tight convergence, with sub-mHa variations among relative energies for the tightest heat-bath tolerances. In hexatriene, the 1^1B_u state qualitatively changes at $\epsilon_{var} = 7 \cdot 10^{-4}$, but the 1^1B_u-1^1A_g gap is well-behaved after this point. As in butadiene, investigations of hexatriene photo dynamics[77–80] place 1^1B_u close in energy to 2^1A_g. At the vertical excitation geometry, SHCI places 2^1A_g above 1^1B_u with a small gap of only 0.08 eV. For the triplet state, the SHCI 1^3B_u-1^1A_g gap agrees well with iFCI at the slightly smaller 6-31G* basis, differing by 0.05 eV (Table 5.6). Once again, the SHCI 1^1B_u-1^1A_g gap differs significantly from experiment,[81] indicating that experimental band maxima do not correspond to vertical excitation energies in hexatriene.

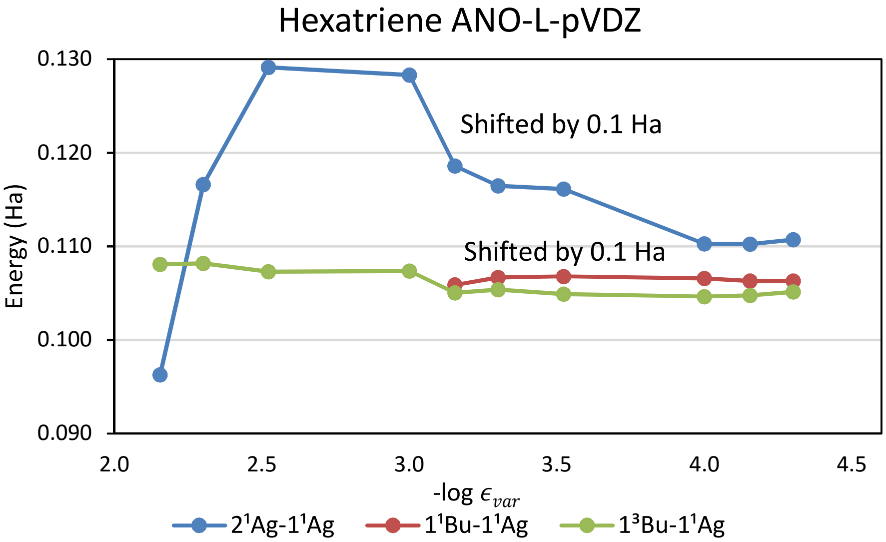


Figure 5.7 Energy gaps versus -log ϵ_{var} for hexatriene/ANO-L-pVDZ. 2^1A_g-1^1A_g and 1^1B_u-1^1A_g values are shifted by 0.1 Ha. Fluctuations in energy gaps settle under 1 mHa by -log ϵ_{var} =4.0, $\epsilon_{var} = 1 \cdot 10^{-4}$ Ha.

Table 5.6 Comparison of SHCI hexatriene/ANO-L-pVDZ energy gaps with literature values. All values in eV.

Gap	f_{PT}(%)	SHCI[a]	CC	iFCI	Exp
2^1A_g-1^1A_g	14.9	5.73	5.72[b]		5.21[e]
1^1B_u-1^1A_g	14.6	5.61	5.30[b]		4.95,[f] 5.13[f]
1^3B_u-1^1A_g	14.7	2.86	2.80[c]	2.81[d]	2.61[f]

[a] E_{SHCI} at $\epsilon_{var} = 5 \cdot 10^{-5}$ Ha
[b] CR-EOMCC(2,3)D/TZVP from reference 38
[c] CCSD(T)/6-31G* from reference 9
[d] iFCI/6-31G* result from reference 9
[e] Raman scattering results from reference 82
[f] Electron impact band maximum from reference 81

5.7 Conclusion

SHCI represents an important step forward for SCI+PT methods, enabling computations of FCI-quality energies in some of the largest systems to date. In order to determine convergence for these polyatomics, a new metric for evaluating SCI+PT energies was introduced. This metric, f_{PT}, was empirically shown to be a reliable indicator of convergence in energy gaps, but not absolute energies. SHCI therefore can reach systems as large as butadiene and hexatriene, where CI spaces of 10^7-10^8 determinants were used to represent FCI spaces of 10^{35} and 10^{38}, respectively, and still produce excitation energies to chemical accuracy.

SHCI achieves this success by treating all valence electrons on an equal footing, and recovering most of the correlation energy in a deterministic fashion. These properties give it advantages over DMRG, which relies on structuring the correlation into an effectively 1-dimensional problem, and over FCIQMC, which relies heavily on stochastic sampling. While the computational scaling of SHCI is still exponential, its ability to systematically reach FCI-level energies (with decreasing ϵ_{var}) means that ground states and excited states in polyatomics with up to ~32 electrons can now be studied with sub-mHa accuracy.

5.8 References

1. Knowles, P. J. & Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **111,** 315–321 (1984).
2. Olsen, J., Roos, B. O., Jørgensen, P. & Jensen, H. J. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* **89,** 2185 (1988).
3. Olsen, J., Jørgensen, P. & Simons, J. Passing the one-billion limit in full configuration-interaction (FCI) calculations. *Chem. Phys. Lett.* **169,** 463–472 (1990).
4. Olsen, J., Christiansen, O., Koch, H. & Jo/rgensen, P. Surprising cases of divergent behavior in Moller–Plesset perturbation theory. *J. Chem. Phys.* **105,** 5082–5090 (1996).
5. Rossi, E., Bendazzoli, G. L., Evangelisti, S. & Maynau, D. A full-configuration benchmark for the N_2 molecule. *Chem. Phys. Lett.* **310,** 530–536 (1999).
6. Dutta, A. & Sherrill, C. D. Full configuration interaction potential energy curves for breaking bonds to hydrogen: An assessment of single-reference correlation methods. *J. Chem. Phys.* **118,** 1610–1619 (2003).
7. Gan, Z., Grant, D. J., Harrison, R. J. & Dixon, D. A. The lowest energy states of the group-IIIA–group-VA heteronuclear diatomics: BN, BP, AlN, and AlP from full configuration interaction calculations. *J. Chem. Phys.* **125,** 124311 (2006).
8. Zimmerman, P. M. Incremental full configuration interaction. *J. Chem. Phys.* **146,** 104102 (2017).

9. Zimmerman, P. M. Singlet–Triplet Gaps through Incremental Full Configuration Interaction. *J. Phys. Chem. A* **121,** 4712–4720 (2017).
10. Zimmerman, P. M. Strong correlation in incremental full configuration interaction. *J. Chem. Phys.* **146,** 224104 (2017).
11. White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B* **48,** 10345–10356 (1993).
12. White, S. R. & Martin, R. L. Ab initio quantum chemistry using the density matrix renormalization group. *J. Chem. Phys.* **110,** 4127–4130 (1999).
13. Chan, G. K.-L. & Head-Gordon, M. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *J. Chem. Phys.* **116,** 4462–4476 (2002).
14. Chan, G. K.-L. & Sharma, S. The Density Matrix Renormalization Group in Quantum Chemistry. *Annu. Rev. Phys. Chem.* **62,** 465–481 (2011).
15. Olivares-Amaya, R. *et al.* The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.* **142,** 34102 (2015).
16. Booth, G. H., Thom, A. J. W. & Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **131,** 54106 (2009).
17. Cleland, D., Booth, G. H. & Alavi, A. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.* **132,** (2010).
18. Chien, A. D. & Zimmerman, P. M. Iterative submatrix diagonalisation for large configuration interaction problems. *Mol. Phys.* **8976,** (2017).
19. Huron, B., Malrieu, J. P. & Rancurel, P. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *J. Chem. Phys.* **58,** 5745–5759 (1973).
20. Buenker, R. J. & Peyerimhoff, S. D. Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theor. Chim. Acta* **35,** 33–58 (1974).
21. Buenker, R. J., Peyerimhoff, S. D. & Butscher, W. Applicability of the multi-reference double-excitation CI (MRD-CI) method to the calculation of electronic wavefunctions and comparison with related techniques. *Mol. Phys.* **35,** 771–791 (1978).
22. Evangelisti, S., Daudey, J. P. & Malrieu, J. P. Convergence of an improved CIPSI algorithm. *Chem. Phys.* **75,** 91–102 (1983).
23. Harrison, R. J. Approximating full configuration interaction with selected configuration interaction and perturbation theory. *J. Chem. Phys.* **94,** 5021 (1991).
24. Ben Amor, N., Bessac, F., Hoyau, S. & Maynau, D. Direct selected multireference configuration interaction calculations for large systems using localized orbitals. *J. Chem. Phys.* **135,** (2011).
25. Tubman, N. M., Lee, J., Takeshita, T. Y., Head-Gordon, M. & Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **145,** 44112 (2016).
26. Holmes, A. A., Tubman, N. M. & Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **12,** 3674–3680 (2016).
27. Sharma, S., Holmes, A. A., Jeanmairet, G., Alavi, A. & Umrigar, C. J. Semistochastic Heat-

Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory. *J. Chem. Theory Comput.* **13,** 1595–1604 (2017).

28. Sherrill, C. D., Leininger, M. L., Van Huis, T. J. & Schaefer, H. F. Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (TZ2P) basis. *J. Chem. Phys.* **108,** 1040 (1998).
29. Theis, D., Ivanic, J., Windus, T. L. & Ruedenberg, K. The transition from the open minimum to the ring minimum on the ground state and on the lowest excited state of like symmetry in ozone: A configuration interaction study. *J. Chem. Phys.* **144,** 104304 (2016).
30. Tavan, P. & Schulten, K. Electronic excitations in finite and infinite polyenes. *Phys. Rev. B* **36,** 4337–4358 (1987).
31. Watts, J. D., Gwaltney, S. R. & Bartlett, R. J. Coupled-cluster calculations of the excitation energies of ethylene, butadiene, and cyclopentadiene. *J. Chem. Phys.* **105,** 6979–6988 (1996).
32. Starcke, J. H., Wormit, M., Schirmer, J. & Dreuw, A. How much double excitation character do the lowest excited states of linear polyenes have? *Chem. Phys.* **329,** 39–49 (2006).
33. Li, X. & Paldus, J. Size dependence of the X^1A_g --> 1^1B_u excitation energy in linear polyenes. *Int. J. Quantum Chem.* **74,** 177–192 (1999).
34. Nakayama, K., Nakano, H. & Hirao, K. Theoretical study of the π-π* excited states of linear polyenes: The energy gap between $1^1B_u^+$ and $2^1A_g^-$ states and their character. *Int. J. Quantum Chem.* **66,** 157–175 (1998).
35. Mazur, G. & Włodarczyk, R. Application of the dressed time-dependent density functional theory for the excited states of linear polyenes. *J. Comput. Chem.* **30,** 811–817 (2009).
36. Schmidt, M. & Tavan, P. Electronic excitations in long polyenes revisited. *J. Chem. Phys.* **136,** 124309 (2012).
37. Schreiber, M., Silva-Junior, M. R., Sauer, S. P. A. & Thiel, W. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* **128,** 1–25 (2008).
38. Piecuch, P., Hansen, J. A. & Ajala, A. O. Benchmarking the completely renormalised equation-of-motion coupled-cluster approaches for vertical excitation energies. *Mol. Phys.* **113,** 3085–3127 (2015).
39. Widmark, P.-O., Malmqvist, P.-Å. & Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor. Chim. Acta* **77,** 291–306 (1990).
40. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90,** 1007–1023 (1989).
41. Kendall, R. a, Dunning, T. H. & Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **96,** 6796–6806 (1992).
42. Krylov, A. I. & Gill, P. M. W. Q-Chem: An engine for innovation. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **3,** 317–326 (2013).
43. Manohar, P. U. & Krylov, A. I. A noniterative perturbative triples correction for the spin-flipping and spin-conserving equation-of-motion coupled-cluster methods with

single and double substitutions. *J. Chem. Phys.* **129,** 194105 (2008).

44. Włoch, M., Lodriguito, M. D., Piecuch, P. & Gour, J. R. Two new classes of non-iterative coupled-cluster methods derived from the method of moments of coupled-cluster equations. *Mol. Phys.* **104,** 2149–2172 (2006).
45. Schaefer, H. F. Methylene: A Paradigm for Computational Quantum Chemistry. *Science.* **231,** 1100–1107 (1986).
46. Bauschlicher, C. W. & Taylor, P. R. A full CI treatment of the $^1A_1$–$^3B_1$ separation in methylene. *J. Chem. Phys.* **85,** 6510 (1986).
47. Sherrill, C. D., Van Huis, T. J., Yamaguchi, Y. & Schaefer, H. F. Full configuration interaction benchmarks for the states of methylene. *J. Mol. Struct. THEOCHEM* **400,** 139–156 (1997).
48. Zimmerman, P. M., Toulouse, J., Zhang, Z., Musgrave, C. B. & Umrigar, C. J. Excited states of methylene from quantum Monte Carlo. *J. Chem. Phys.* **131,** 124103 (2009).
49. Slipchenko, L. V. & Krylov, A. I. Singlet-triplet gaps in diradicals by the spin-flip approach: A benchmark study. *J. Chem. Phys.* **117,** 4694–4708 (2002).
50. Shao, Y., Head-Gordon, M. & Krylov, A. I. The spin–flip approach within time-dependent density functional theory: Theory and applications to diradicals. *J. Chem. Phys.* **118,** 4807 (2003).
51. Chien, A. D. & Zimmerman, P. M. Recovering dynamic correlation in spin flip configuration interaction through a difference dedicated approach. *J. Chem. Phys.* **146,** 14103 (2017).
52. Jensen, P. & Bunker, P. R. The potential surface and stretching frequencies of X^3B_1 methylene (CH_2) determined from experiment using the Morse oscillator-rigid bender internal dynamics Hamiltonian. *J. Chem. Phys.* **89,** 1327–1332 (1988).
53. Alijah, A. & Duxbury, G. Renner-Teller and spin-orbit interactions between the $^1A_1$, $^1B_1$ and $^3B_1$ states of CH_2. *Mol. Phys.* **70,** 605–622 (1990).
54. Davidson, E. R. The Spatial Extent of the V State of Ethylene and Its Relation to Dynamic Correlation in the Cope Rearrangement. *J. Phys. Chem.* **100,** 6161–6166 (1996).
55. Müller, T., Dallos, M. & Lischka, H. The ethylene 1^1B_{1u} V state revisited. *J. Chem. Phys.* **110,** 7176–7184 (1999).
56. Angeli, C. An analysis of the dynamic σ polarization in the V state of ethene. *Int. J. Quantum Chem.* **65,** NA-NA (2010).
57. Daday, C., Smart, S., Booth, G. H., Alavi, A. & Filippi, C. Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question. *J. Chem. Theory Comput.* **8,** 4441–4451 (2012).
58. Mulliken, R. S. The excited states of ethylene. *J. Chem. Phys.* **66,** 2448–2451 (1977).
59. Moore, J. H. & Doering, J. P. Ion Impact Spectroscopy: Inelastic Scattering of 150–500-eV H^+ and H_2^+ from N_2 , CO, C_2H_2 , and C_2H_4. *J. Chem. Phys.* **52,** 1692–1699 (1970).
60. Moore, J. H. Investigation of the low energy singlet-triplet and singlet-singlet transitions in ethylene derivatives by ion impact. *J. Phys. Chem.* **76,** 1130–1133 (1972).
61. Van Veen, E. H. Low-energy electron-impact spectroscopy on ethylene. *Chem. Phys. Lett.* **41,** 540–543 (1976).
62. Andersen, S. O., Halberstadt, M. L. & Borgford-Parnell, N. Stratospheric ozone, global warming, and the principle of unintended consequences—An ongoing science and

policy success story. *J. Air Waste Manage. Assoc.* **63,** 607–647 (2013).

63. Hay, P. J. & Goodard, W. A. Theoretical results for the excited states of ozone. *Chem. Phys. Lett.* **14,** 46–48 (1972).
64. Lee, T. J. On the energy separation between the open and cyclic forms of ozone. *Chem. Phys. Lett.* **169,** 529–533 (1990).
65. Qu, Z.-W., Zhu, H. & Schinke, R. Infrared spectrum of cyclic ozone: A theoretical investigation. *J. Chem. Phys.* **123,** 204324 (2005).
66. Xantheas, S., Elbert, S. T. & Ruedenberg, K. An intersection seam between the ground state of ozone and an excited state of like symmetry a). *J. Chem. Phys.* **93,** 7519–7521 (1990).
67. Xantheas, S. S., Atchity, G. J., Elbert, S. T. & Ruedenberg, K. Potential energy surfaces of ozone. I. *J. Chem. Phys.* **94,** 8054–8069 (1991).
68. Atchity, G. J. & Ruedenberg, K. Global potential energy surfaces for the lowest two $1A'$states of ozone. *Theo. Chem. Acct. Theory, Comput. Model. (Theoretica Chim. Acta)* **96,** 176–194 (1997).
69. Atchity, G. J., Ruedenberg, K. & Nanayakkara, A. The intersection seam between the $1^1A'$ and $2^1A'$ states of ozone. *Theo. Chem. Acct. Theory, Comput. Model. (Theoretica Chim. Acta)* **96,** 195–204 (1997).
70. Komainda, A., Lefrancois, D., Dreuw, A. & Köppel, H. Theoretical study of the initial non-radiative $1B_u \rightarrow 2A_g$ transition in the fluorescence quenching of s-trans-butadiene: Electronic structure methods and quantum dynamics. *Chem. Phys.* **482,** 27–38 (2017).
71. Fuß, W., Schmid, W. E. & Trushin, S. A. Ultrafast electronic relaxation of s-trans-butadiene. *Chem. Phys. Lett.* **342,** 91–98 (2001).
72. Watson, M. A. & Chan, G. K.-L. Excited States of Butadiene to Chemical Accuracy: Reconciling Theory and Experiment. *J. Chem. Theory Comput.* **8,** 4013–4018 (2012).
73. Mosher, O. A., Flicker, W. M. & Kuppermann, A. Electronic spectroscopy of s-trans 1,3-butadiene by electron impact. *J. Chem. Phys.* **59,** 6502–6511 (1973).
74. McDiarmid, R. On the ultraviolet spectrum of trans-1,3-butadiene. *J. Chem. Phys.* **64,** 514–521 (1976).
75. Doering, J. P. & McDiarmid, R. Electron impact study of the energy levels of trans-1,3-butadiene: II. Detailed analysis of valence and Rydberg transitions. *J. Chem. Phys.* **73,** 3617–3624 (1980).
76. Mosher, O. A., Flicker, W. M. & Kuppermann, A. Triplet states in 1,3-butadiene. *Chem. Phys. Lett.* **19,** 332–333 (1973).
77. Hayden, C. C. & Chandler, D. W. Femtosecond Time-Delayed Photoionization Studies of Ultrafast Internal Conversion in 1,3,5-Hexatriene. *J. Phys. Chem.* **99,** 7897–7903 (1995).
78. Cyr, D. R. & Hayden, C. C. Femtosecond time-resolved photoionization and photoelectron spectroscopy studies of ultrafast internal conversion in 1,3,5-hexatriene. *J. Chem. Phys.* **104,** 771–774 (1996).
79. Ohta, K., Naitoh, Y., Tominaga, K., Hirota, N. & Yoshihara, K. Femtosecond Transient Absorption Studies of trans - and cis -1,3,5-Hexatriene in Solution. *J. Phys. Chem. A* **102,** 35–44 (1998).
80. Komainda, A., Lyskov, I., Marian, C. M. & Köppel, H. Ab Initio Benchmark Study of Nonadiabatic S_1 –S_2 Photodynamics of cis - and trans -Hexatriene. *J. Phys. Chem. A*

120, 6541–6556 (2016).
81. Flicker, W. M., Mosher, O. A. & Kuppermann, A. Low energy, variable angle electron-impact excitation of 1,3,5-hexatriene. *Chem. Phys. Lett.* **45,** 492–497 (1977).
82. Fujii, T., Kamata, A., Shimizu, M., Adachi, Y. & Maeda, S. Two-photon absorption study of 1,3,5-hexatriene by CARS and CSRS. *Chem. Phys. Lett.* **115,** 369–372 (1985).

Chapter 6: Final Remarks

By harnessing light, researchers gain access to a wide regime of photochemical transformations and processes that can be used in technologies that benefit human civilization. The development and optimization of these technologies requires an in-depth understanding of excited state dynamics, which can be provided by computational chemistry methods. However, the widespread adoption of computational chemistry is currently obstructed by two major hurdles. First, the wide variety of state characters are difficult to treat equally within a single truncated CI method, which can lead to excited state energy gaps that are qualitatively incorrect. Second, CI methods that treat all states equally are typically very expensive, utilizing vast amounts of computational resources and necessitating specialized "clusters" of computers. The efficient treatment of both of these problems are the main areas to which this dissertation has contributed.

The initial Chapters focused on treating states of disparate character equally with SF-CI methods, which utilize a high-spin reference instead of the typical HF reference, and are particularly well-suited for the multiexciton TT states in singlet fission. Chapter 2's investigation of a quinoidal bithiophene with RAS(h,p)-SF determined that generation of independent triplets was unfeasible given experimental conditions, leading to the hypothesis of a long-lived 2^1A_g state. The single exciton states of the quinoidal bithiophene were only qualitatively treated by RAS(h,p)-SF, necessitating the use of XMS-CASPT2 methods to obtain a holistic energetic picture of the excited state dynamics. Extensions to RAS(h,p)-SF were developed in Chapter 3 in an effort to increase the range of states that RAS-SF methods could be quantitatively applied to. The resulting method, RAS(S,2h,2p)-SF, demonstrated the importance of extended descriptions of the ^{5}(TT) state, placing ^{5}(TT) below ^{1}(TT) in a tetracene dimer, in direct opposition to predictions obtained with the lower RAS(h,p)-SF method. However, single exciton states remained qualitatively described.

These results illuminate the difficulties inherent in solving the problem of unbalanced references. While SF-CI methods treat excited and ground states on a more balanced footing, it still fails to treat *all* states equally, returning only qualitative single exciton wave functions. Straightforward treatment of this problem by expanding the CI space with various classes of excitations, as in RAS(S,2h,2p)-SF, was unable to obtain quantitative energies. Alternatives, such as MCSCF, require one to tailor chosen reference configurations to the states and systems of interest. An open question in the field then, is how one can generate balanced references in an automated manner that is not too costly? If such a method existed, quantitative energy gaps for states of many different characters could be obtained at truncated CI levels, greatly increasing the system sizes to which CI could be applied for excited state studies.

In lieu of balanced references, one can develop strategies for improving the efficiency of FCI, which does not depend on the reference choice. The ISD algorithm presented in Chapter 4 took advantage of natural orbitals to divide the Hamiltonian into a series of increasingly accurate submatrices to be solved iteratively. A well-behaved error metric, the energy difference between iterations, allowed one to stop iterating once errors fell below chemical accuracy, resulting in about an order of magnitude reduction in CI space. However, much larger reductions are necessary to reach a broader range of molecules. Chapter 5 applied the SHCI algorithm towards excited states, reducing the CI space by many orders of magnitude, treating a FCI space of 10^{38} with only 10^{8} determinants. This allowed the calculation of valence energy gaps on butadiene and hexatriene, two systems typically thought of as outside of the reach of FCI-level methods.

To push FCI-level calculations to ever larger systems, attributes of the various methods that approximate FCI will likely need to be combined. For instance, SHCI already combines CI, PT, and stochastic methodology to great effect. The ISD and SHCI strategies are some of the more flexible FCI approximations, being applicable to excited states while avoiding strong dependencies on inherent Hamiltonian structure (as in DMRG) or stochastic sampling (as in FCIQMC). However, there are some properties that SHCI could benefit from, such as the low basis set dependence of the DMC methodology. Especially important to enlarging future FCI-level calculations will be reducing the exponential scaling of FCI. So far, only iFCI has managed to reduce the FCI problem to polynomial scaling, although it is currently only applicable to ground states. Although SHCI has proven very powerful already, implementing the advantages inherent in other FCI approximations would increase its flexibility and enable it to be applied towards much larger systems.

There are a number of ways in which once could combine or refine the methods discussed in this dissertation. For instance, the generally applicable strategies of Chapters 4 and 5 could be applied to the RAS-SF methodology, to reduce the cost of including more excitation classes. This would allow for studies of singlet fission systems that go beyond minimal model systems, which only approximate the true physical picture. The ISD strategy could also be refined with a dynamic $\eta^{(i)}$ series that changes with the system, as different systems may have different natural orbital occupation distributions. Variants of HCI that focus on energy differences, as in RAS(S,2h,2p)-SF and DDCI, could be developed, further reducing the necessary CI space for quantitative excitation energies. Further investigation into the stability of the f_{PT} metric is also warranted, as only a few test cases were used to establish its practicality.

This dissertation has extended the reach of CI methods for excited state quantities. However, much work remains to be done in the development of perfectly balanced references and FCI approximations. RAS(S,2h,2p)-SF enables one to deeply inspect the TT states of singlet fission, but shows indications that the high-spin references are unbalanced with respect to ground and single exciton states. With SHCI, FCI benchmarks are now possible for systems with about 32 electrons. Future developments of more balanced references or more efficient FCI approximations are greatly anticipated, as either would pave the way for computational chemistry to play a commanding role in the development of excited state chemistries.

Appendix A: Supporting Data for Chapter 2

A.1 Computational Details/Data

A.1.1 XMS-CASPT2 Geometries and Single Points

The first step in these calculations is to define an active space, as a CASSCF wave function and orbitals are necessary inputs for the XMS-CASPT2 computation. We chose an active space of 10 electrons in 8 π-orbitals, (10,8), in part due to the findings of the previous study.[1] Analysis of the singlet wave functions indicated that the HOMO-4 π orbital had significant contributions to the ^1B_u states, which come into play in the transition dipole moment calculations. Orbitals are discussed in terms of the irreducible representations of the C2h symmetry in the following order as in MOLPRO - (A_g, A_u, B_u, B_g).

We targeted five states, two ^1A_g and three ^1B_u, based on which states were available in the range of the experimental pump and probe energies (2.8 eV and 1.4 eV). All states were evenly weighted in the state-averaged CASSCF calculations. For the triplet manifold, we targeted only two states: one ^3B_u and one ^3A_g. XMS-CASPT2 computations utilized MOLPRO's RS2 module. A level shift of 0.2 au was used to avoid the effect of intruder states.

A.1.2 MCTDH Details

The quantum dynamics simulations were conducted using the Multiconfigurational Time Dependent Hartree-Fock (MCTDH) method, an efficient algorithm for solving the time-dependent Schrodinger equation. The vibronic coupling Hamiltonian was truncated at linear expansion, as shown in Equation A1, for four modes (three symmetric tuning modes and one asymmetric coupling mode). We generate the quasi-diabatic representation for the 1^1B_u and 2^1A_g states by minimizing the squared error between the eigenvalues of the 2x2 diabatic potential energy matrix and the state energies from RAS-SF(6,6) calculations (Figure A.1).

The three tuning modes were chosen as Condon-active symmetrical modes modulating the energy gap between 1^1B_u and 2^1A_g. The asymmetric coupling mode was chosen because it corresponded with the non-adiabatic coupling vector between 1^1B_u and 2^1A_g as calculated at the CASSCF(4,4) level. Thus, the model Hamiltonian incorporates a quantitative measure of the coupling between the two states of interest through diabatization along the asymmetric coupling mode.

We note that, along the symmetrical tuning modes, there should be no coupling between the states due to symmetry. Thus, along the tuning modes, a fit is being made to the adiabatic RAS-SF(6,6) potentials. The asymmetrical coupling mode however, breaks symmetry and induces coupling between 1^1B_u and 2^1A_g. Thus, along the coupling mode, diabatization is occurring.

$$\hat{H}_d = T(q) + \begin{bmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{bmatrix} + \sum_{i=1}^{4} \left(\frac{\omega_i}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} q_i^2 + \begin{bmatrix} a_i^{(1)} & b_i \\ b_i & a_i^{(2)} \end{bmatrix} q_i \right)$$

Equation A.1 Form of the linear vibronic coupling Hamiltonian. *T(q)* represents the kinetic energy operator while ε, ω, *a*, and *b*, are fit using the RAS-SF data.

The resulting LVCH shows no interstate coupling for the three symmetric, tuning, modes (9, 48, 50) as expected from symmetry. Parameters for the fit are reported in Table A.1. For each vibrational mode, the residual sum of squares was less than 0.5 eV.

Table A.1 Fitted parameters for the MCTDH simulations. All values are in Hartree.

Mode	Linear Coupling Parameters			Ground State Frequency	Excited State Frequency (w)	MSE for fit
	a(1)	a(2)	b			
9	4.61E-03	1.45E-03	4.52E-04	1.10E-03	3.38E-03	1.21E-05
48	-1.39E-02	-3.18E-03	0.00E+00	7.43E-03	5.61E-03	1.13E-04
50	1.06E-02	5.52E-03	0.00E+00	7.50E-03	2.20E-03	2.96E-04
51	0.00E+00	0.00E+00	1.82E-03	7.60E-03	7.60E-03	2.40E-04

epsilon1 = epsilon2 = 6.90E-03 Ave. MSE: 0.000165
Bath maximum frequency = 1.80E-02
bath coupling strength = 4.05E-04

To account for energy dissipation in this system, a weakly interacting Ohmic bath of 20 harmonic oscillators was included in the model. The cutoff frequency ω_f of the bath was chosen to be ~2.2 times the largest frequency in the four-mode subsystem in order to avoid resonance with the system modes while maintaining a sufficiently long recurrence times (Table A.1).

We also include plots of the expectation values for the vibrational quantum number, <n>, of each mode as a function of time (fs) (Figure A.1).

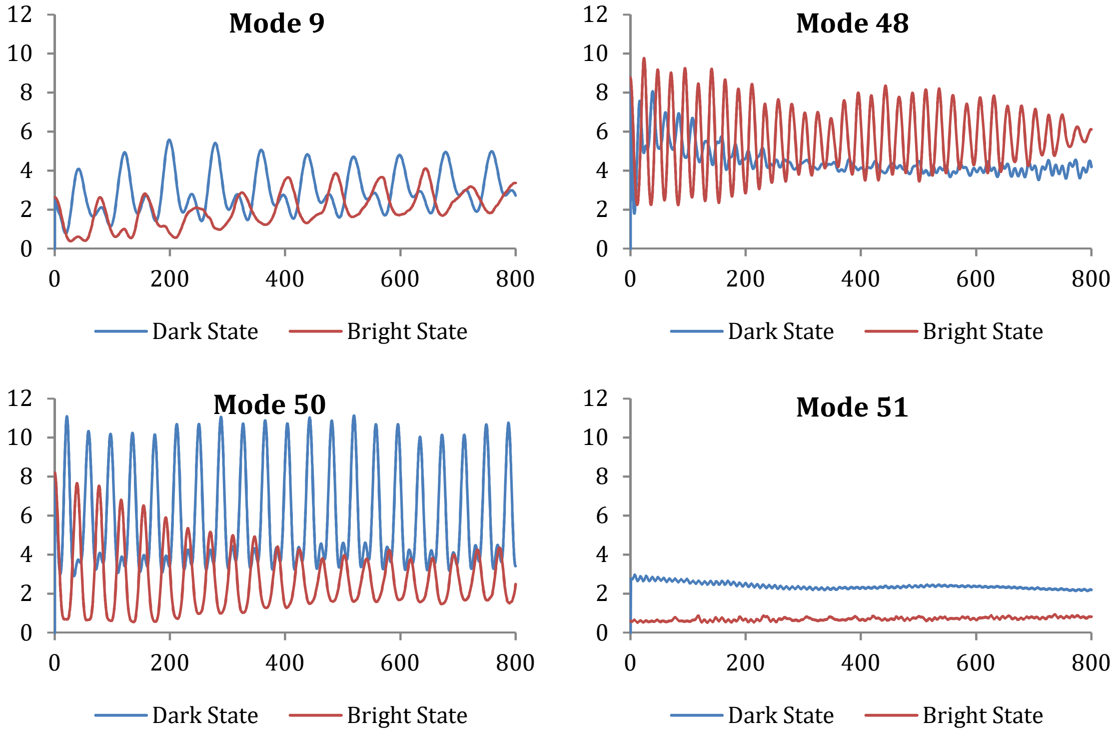


Figure A.1 Expectation values of vibrational quantum number in each mode over time.

A.2 Spin-Orbit Coupling

The spin-orbit couplings (SOC's) between the 2^1A_g and T_2 (1^3A_g) states of QOT2 were computed using CASSCF(10,8) wave functions at the XMS-CASPT2 2^1A_g geometry using MOLPRO. Three states were targeted: two ^1A_g and one ^3A_g state, with the two ^1A_g

states weighted as 0 and 1, to obtain an accurate description of the two states of interest (1^3A_g and 2^1A_g). MOLPRO provides both mean-field and Breit-Pauli spin operators. The value quoted in the paper, 0.0043 cm^{-1}, is from the Breit-Pauli operator. The Breit-Pauli operator was chosen because includes both one and two-electron operators as opposed to the mean field – which approximates the two electron terms.[2] The mean-field operator gives 0.00043 cm^{-1}. We then calculated a Marcus rate for ISC in the high-temperature limit as described by Beljonne et al.[3] This calculation is detailed below.

The golden rule expression for radiationless transitions, shown below, requires the Franck-Condon weighted density of states (FCWD).

$$k_{ISC}^{IF} = \frac{2\pi}{\hbar}\langle {}_{I}^{1}\Psi|H_{SO}|{}_{F}^{3}\Psi\rangle^2 \cdot FCWD$$

By taking the high-temperature limit, we can write FCWD as,

$$FCWD = \frac{1}{\sqrt{4\pi\lambda RT}}\exp\left[-\frac{(\Delta E+\lambda)^2}{4\lambda RT}\right]$$

where λ is the Marcus reorganization energy, R the gas constant, T the temperature (298K), and ΔE the energy difference between the final and initial states. We were unable to obtain the 1^3A_g geometry at the XMS-CAPT2 (10,8) level due to poor convergence of the optimizer. Examining the 1^3A_g wave function indicates small contributions from the highest and lowest LUMO and HOMO in the (10,8) space. So, we instead use an active space of (8,6) to optimize the 1^3A_g geometry, which converged satisfactorily. The resulting values for ISC from 2^1A_g to 1^3A_g are

$$\lambda = 0.15 \text{ eV},\ \Delta E = 0.14 \text{ eV}, RT = 0.0257 \text{ eV, with } \hbar = 6.582\cdot 10^{-16}\text{eV·s}$$

Utilizing 0.0043 cm^{-1} = 5.33x10^{-7} eV, we obtain a rate constant of

$$k_{ISC}^{IF} = \frac{2\pi}{\hbar}(5.33\cdot 10^{-7})^2\cdot\frac{1}{\sqrt{4\pi\cdot 0.150\cdot 0.0257}}\exp\left[-\frac{(0.290)^2}{4\cdot 0.150\cdot 0.0257}\right] = 52.73\text{ s}^{-1}$$

This is equivalent to a time constant of 18.9 ms.

While QOT2's optimized structure is planar in the electronic states of interest, vibrational motion breaks this planarity. To approximate the effect of twisting on the degree of spin-orbit coupling, we calculated SOC's at 15, 30, 45, and 60 degrees rotation about the central C=C bond. These SOC calculations were set up slightly different because of the loss of symmetry due to the C=C twist. We used the same (10,8) active space, using visual inspection to insure that all π orbitals were chosen for the active space. Four states were targeted: the two lowest singlets and the two lowest triplets. Weighting for the CASSCF procedure was set to 0,1 for both singlets and triplets, making it so that we targeted only the second highest state in both the singlet and triplet manifolds, which corresponds to the 2^1A_g and 1^3A_g states of interest. The resulting SOC's are tabulated in Table A.2. They all stay quite low, indicating that ISC will play a negligible role in the dynamics of the 2^1A_g exciton, even with twisting.

Table A.2 Computed SOC for twisted QOT2. The main row to note is the bottom one, which square-roots the sum of the squares of the X , Y, and Z components and indicates that the overall SOC is small. This means the rate will also be small due to the $\langle {}^1_I\Psi | H_{SO} | {}^3_F\Psi \rangle^2$ factor in the rate equation above.

Twist (degrees)	15		30	
Operator	Mean-Field	Breit-Pauli	Mean-Field	Breit-Pauli
$H_{SO,x}$ (cm^{-1})	1.71E-06	1.26E-04	2.25E-04	4.49E-04
$H_{SO,y}$ (cm^{-1})	-3.87E-05	-4.54E-04	-7.19E-04	-1.47E-03
$H_{SO,z}$ (cm^{-1})	-3.34E-04	-4.99E-03	-4.02E-03	-8.15E-03
sqrt(sum^2) (cm-1)	3.36E-04	5.02E-03	4.09E-03	8.29E-03
sqrt(sum^2) (eV)	4.17E-08	6.22E-07	5.07E-07	1.03E-06

	45		60	
Operator	Mean-Field	Breit-Pauli	Mean-Field	Breit-Pauli
$H_{SO,x}$ (cm^{-1})	2.14E-03	2.39E-03	5.14E-03	5.27E-03
$H_{SO,y}$ (cm^{-1})	-1.94E-03	-2.78E-03	-1.77E-02	-1.77E-02
$H_{SO,z}$ (cm^{-1})	-1.31E-02	-1.60E-02	-4.48E-02	-4.58E-02
sqrt(sum^2) (cm^{-1})	1.34E-02	1.65E-02	4.84E-02	4.94E-02
sqrt(sum^2) (eV)	1.66E-06	2.04E-06	6.00E-06	6.12E-06

A.3 Pentacene Dimer Natural Orbitals

A publication by Zimmerman et al.[4] calculated the geometry of an embedded pentacene dimer and found that the excimer relaxes along an inter-pentacence C-C coordinate from 5.6 Å to 5.3 Å separation. Thus, the occupancy numbers and natural orbitals of the ^{1}(TT) state were calculated at the 5.3 Å geometry. RAS-SF(4,4) with a quintet reference and 6-31G* was used.

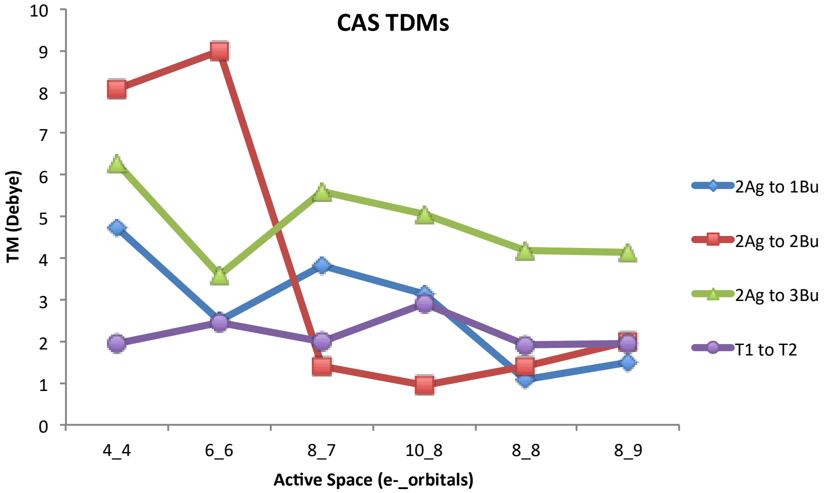
A.4 Localized Orbitals

Localized orbitals were obtained via Pipek-Mezey localization[5] of the four frontier (HOMO-1 to LUMO+1) ROHF quintet molecular orbitals.

A.5 Constrained Hartree-Fock

We used constrained HF in Q-Chem to calculate an approximate geometry for two separated triplets in the 6-31G* basis. An αα triplet was placed on the left half of QOT2 and a ββ triplet on the right.

A.6 TDM Stability



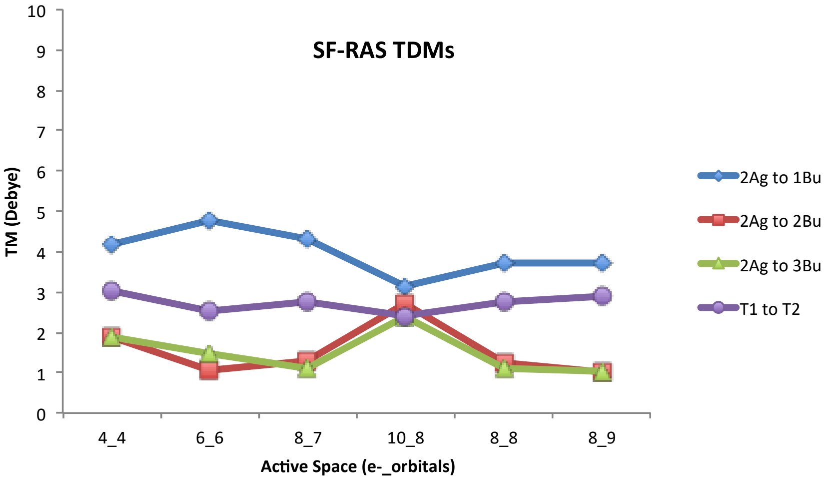


Figure A.2 TDM calculations for CASSCF and RAS-SF methods. These preliminary TDM calculations were run at geometries obtained at the CASSCF(10,8) and RAS-SF(10,8) levels . Although the major variation in CASSCF TDMs occurs at the lower active spaces, the RAS-SF TDMs have greater overall stability due to inclusion of increased σ-π dynamic correlation. Analysis of the wave functions indicate that the HOMO-4 orbital is integral to describing the 2 and 3^1B_u states suggesting the (10,8) active space is most appropriate for TDM calculations.

A.7 3^1A_g Energy Estimate

The energy of the 3^1A_g state was estimated using a XMS-CASPT2 calculation at the 1^1A_g geometry with the (10,8) active space. Six states were computed: three ^1A_g and three ^1B_u. Single points at this geometry showed that the 3^1A_g is located 1.26 eV above 1^1B_u. 1^1B_u

itself is located is located 2.69 eV above 1^1A_g. Given the pump beam's energy (445 nm, 2.78 eV), it is therefore unlikely that the 3^1A_g surface is accessible.

A.8 References

[1] O. Varnavski, N. Abeyasinghe, J. Aragó, J.J. Serrano-Pérez, E. Ortí, J.T. López Navarrete, K. Takimiya, D. Casanova, J. Casado, and T. Goodson, J. Phys. Chem. Lett. 1375 (2015).
[2] C.M. Marian, Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 187 (2012).
[3] D. Beljonne, Z. Shuai, G. Pourtois, and J.L. Bredas, J. Phys. Chem. A **105**, 3899 (2001).
[4] P.M. Zimmerman, F. Bell, D. Casanova, and M. Head-Gordon, J. Am. Chem. Soc. **133**, 19944 (2011).
[5] J. Pipek and P.G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

Appendix B: Supporting Data for Chapter 3

B.1 Methylene

Table B.1 Methylene absolute energies and excitation energy gaps without adiabatic corrections.

	(2,2)			RAS(S,2h,2p) -SF (4,4)	FCI
	RAS(h,p)-SF	RAS(S)-SF	RAS(S,2h,2p)-SF		
Absolute Energies (H)					
$\tilde{X}$	-38.932976	-38.950117	-38.957880	-39.012221	-39.066738
$\tilde{a}$	-38.910178	-38.932022	-38.937837	-38.992231	-39.048984
$\tilde{b}$	-38.870739	-38.886496	-38.901123	-38.955164	-39.010059
$\tilde{c}$	-38.814431	-38.815252	-38.858711	-38.914948	-38.968471
Adiabatic Gap (eV)					
$\tilde{X} \rightarrow \tilde{a}$	0.620	0.492	0.545	0.544	0.482
$\tilde{X} \rightarrow \tilde{b}$	1.693	1.731	1.544	1.552	1.542
$\tilde{X} \rightarrow \tilde{c}$	3.226	3.670	2.698	2.647	2.674

Adiabatic corrections were neglected for the reported RAS(S,2h,2p)-SF calculations on methylene (Table 3.1) so that comparisons between RAS-SF methods would be fair, since RAS(h,p)-SF and RAS(S)-SF do not have corresponding corrections. Correlation energy contributions from *2h-2p* excitations are given in Table B.2. The corrected RAS(S,2h,2p)-SF values in Table B.3 are derived from the difference between the two states of interest. For instance, the correction for the (2,2) $\tilde{X} \rightarrow \tilde{a}$ gap would be (-1.396- -1.340) = -0.056 eV. The corrections slightly improve RAS(S,2h,2p)-SF values compared to FCI results.

Table B.2 Correlation energy contributions (in eV) from *2h-2p* excitations in methylene.

	(2,2)	(4,4)
$\tilde{X}$	-1.340	-0.270
$\tilde{a}$	-1.396	-0.290
$\tilde{b}$	-1.331	-0.264
$\tilde{c}$	-1.298	-0.242

Table B.3 Adiabatic gaps with corrected RAS(S,2h,2p)-SF values. RAS(h,p)/(S)-SF and FCI values are reproduced for comparison purposes only and are the same as in Table B.1.

Adiabatic Gap (eV)	(2,2)			Corrected RAS(S,2h,2p)-SF (4,4)	FCI
	RAS(h,p)-SF	RAS(S)-SF	Corrected RAS(S,2h,2p)-SF		
$\tilde{X} \rightarrow \tilde{a}$	0.620	0.492	0.489	0.524	0.482
$\tilde{X} \rightarrow \tilde{b}$	1.693	1.731	1.553	1.558	1.542
$\tilde{X} \rightarrow \tilde{c}$	3.226	3.670	2.740	2.675	2.674

B.2 Trimethyleneethane (TME)

Adiabatic corrections were neglected for the reported RAS(S,2h,2p)-SF calculations on TME (Table 3.2 and Figure 3.3). This was done so that comparisons between RAS-SF methods would be fair, since RAS(h,p)-SF and RAS(S)-SF do not have corresponding corrections. The corrections are tabulated in Table B.4. The (2,2) calculations have a maximum correction of 0.02 eV, while the (6,6) corrections have a maximum of 0.005 eV. The corrections do not change the conclusions reached by examination of (2,2) and (6,6) potentials.

Table B.4 TME correlation energy from excluded *2h-2p* excitations and resulting adiabatic corrections.

	(2,2)			(6,6)		
	Correlation Energy (eV)		Correction (eV)	Correlation Energy (eV)		Correction (eV)
Torsion (degrees)	S_0	T_1		S_0	T_1	
0	-21.1609	-21.1411	0.0199	-14.3712	-14.3714	-0.0003
10	-21.1509	-21.1321	0.0188	-14.3614	-14.3603	0.0011
20	-21.1101	-21.0940	0.0161	-14.3189	-14.3181	0.0008
30	-21.0529	-21.0390	0.0139	-14.2596	-14.2615	-0.0019
40	-20.9928	-20.9830	0.0098	-14.2014	-14.2052	-0.0038
45	-20.9765	-20.9669	0.0095	-14.1750	-14.1790	-0.0041

50	-20.9726	-20.9634	0.0093	-14.1513	-14.1556	-0.0044
60	-20.9623	-20.9555	0.0068	-14.1116	-14.1162	-0.0046
70	-20.9473	-20.9435	0.0038	-14.0822	-14.0873	-0.0052
80	-20.9000	-20.9005	-0.0005	-14.0637	-14.0691	-0.0054
90	-20.9133	-20.9141	-0.0008	-14.0560	-14.0615	-0.0054

B.3 Tetracene

B.3.1 Monomer

To select the basis for use in dimer calculations, a series of monomer calculations were run with differing basis sets (Table B.5). The basis that best balances cost versus accuracy, cc-pVTZ with f functions removed for carbon, and cc-pVDZ for hydrogen (C:cc-pVTZ-f/H:cc-pVDZ), was chosen for dimer calculations.

Table B.5 RAS-SF (2,2) calculations on a tetracene monomer with varying basis sets. The S_0 values are absolute energies in Hartree, whereas the S_1, S_2, T_1, and T_2 values are vertical excitation energies, in eV. The first two columns use cc-pVTZ/pVDZ basis for all atoms. The following columns use shorthand to designate basis choice for <u>C</u>arbon and <u>H</u>ydrogen, where tz = cc-pVTZ, dz = cc-pVDZ, tz-f = cc-pVTZ minus f functions, etc.

	RAS(h,p)-SF					
	cc-pVTZ	cc-pVDZ	C:tz/H:dz	C:tz-f/H:dz	C:tz-f/H:tz	C:tz-f/H:tz-d
S_0 (H)	-688.85183	-688.69038	-688.83359	-688.80396	-688.82153	-688.81935
S_1 (eV)	4.022	4.113	4.022	4.025	4.023	4.025
S_2 (eV)	4.893	4.942	4.895	4.890	4.888	4.888
T_1 (eV)	1.935	1.948	1.936	1.932	1.931	1.931
T_2 (eV)	3.563	3.574	3.565	3.559	3.557	3.557
	RAS(S)-SF					
	cc-pVTZ	cc-pVDZ	C:tz/H:dz	C:tz-f/H:dz	C:tz-f/H:tz	C:tz-f/H:tz-d
S_0 (H)	-688.90533	-688.74349	-688.88710	-688.85752	-688.87510	-688.87293
S_1 (eV)	3.702	3.797	3.703	3.712	3.710	3.712
S_2 (eV)	4.649	4.674	4.649	4.639	-	-
T_1 (eV)	1.858	1.881	1.858	1.857	1.856	1.856
T_2 (eV)	3.956	3.965	3.957	3.951	3.950	3.950
	RAS(S,2h,2p)-SF					
	cc-pVTZ	cc-pVDZ	C:tz/H:dz	C:tz-f/H:dz	C:tz-f/H:tz	C:tz-f/H:tz-d
S_0 (H)	-688.91303	-688.75091	-688.89479	-688.86512	-688.88271	-688.88054
S_1 (eV)	3.506	3.603	3.506	3.516	3.514	3.516
S_2 (eV)	4.790	4.825	4.803	4.791	4.784	4.784
T_1 (eV)	2.033	2.049	2.032	2.028	2.028	2.028

T_2 (eV)	4.135	4.136	4.134	4.127	4.126	4.126
# basis fxns	708	312	600	474	582	522

B.3.2 Dimer

All dimer calculations utilize the C:cc-pVTZ-f/H:cc-pVDZ basis chosen via inspection of Table B7. A number of virtual orbitals were also frozen. The number of virtual orbitals was chosen based on when RAS(S)-SF vertical gaps were converged to 0.001 eV for the first five singlet states (Table B.6) – which was at 300 frozen virtuals out of 826 total virtual orbitals.

Table B.6 RAS(S)-SF tetracene dimer energies with varying frozen virtual orbitals.

	RAS(S)-SF				
# frozen virtuals	700	500	400	300	250
S_0 (H)	-1377.67552	-1377.69067	-1377.69274	-1377.69324	-1377.69361
S_1 (eV)	3.698	3.773	3.772	3.773	3.773
S_2 (eV)	3.899	3.819	3.821	3.821	3.820
S_3 (eV)	4.033	3.928	3.920	3.921	3.922
S_4 (eV)	4.254	4.190	4.189	4.189	4.190
S_5 (eV)	4.958	4.905	4.904	4.905	4.905

Appendix C: Supporting Data for Chapter 4

C.1 Comparison with Other Selected CI Methods

Here is a comparison of ISD to two other selected CI methods, CIPSI and HCI, which endeavor to obtain FCI quality energies. Comparisons between calculations converging energies to within 1 mHa of the FCI energy are presented in Table C.1, which shows that ISD's timings are better than CIPSI, while HCI is much quicker than either. The combination of ISD+HCI surprisingly performs slightly worse than plain HCI. Looking closer, one sees that the final ISD+HCI iteration, which considered all orbitals, only took 15.4 sec, significantly quicker than the plain HCI calculation considering all orbitals. However, the cumulative cost of getting to the final iteration abolishes the savings of the final iteration. We see that ISD+HCI timings are the result of an interplay between the cost of earlier iterations and the savings garnered in later ones. The true advantage of combining HCI and ISD lies in its ability to overcome the memory bottlenecks that come with larger systems, as in the N_2 dissociation in Chapter 4.

Table C.1 Cost comparison between ISD, CIPSI, and HCI calculations (28 cores) converged to within 1 mHa of the FCI energy for OH cation in the cc-pVTZ basis.

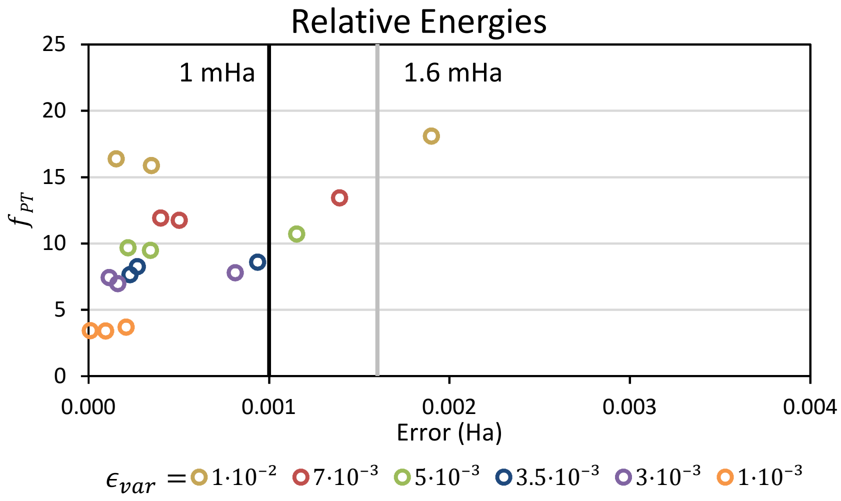
	ISD	CIPSI[a]	HCI	ISD+HCI
Timing (sec)	912.5	1609.8	24.3	37.3

[a] Timings from ORCA 4.0's CIPSI algorithm

Appendix D: Supporting Data for Chapter 5

D.1 Methylene

For methylene/aug-cc-pVQZ, steps of $0.05 \cdot 10^{-3}$ Ha were used from $\epsilon_{var} = 7 \cdot 10^{-3}$ to $1 \cdot 10^{-3}$ Ha to allow for accurate estimates of f_{PT}'s necessary for convergence. Figure D.1 shows f_{PT} behavior as ϵ_{var} decreases. As reported in Table 5.1, the relative energies are converged to chemical accuracy with an average f_{PT} = 12.4% and to sub-mHa with an average f_{PT} = 8.2%. For absolute energies, chemical accuracy is obtained with an average f_{PT} = 10.0% and sub-mHa errors with an average f_{PT} = 6.6%.



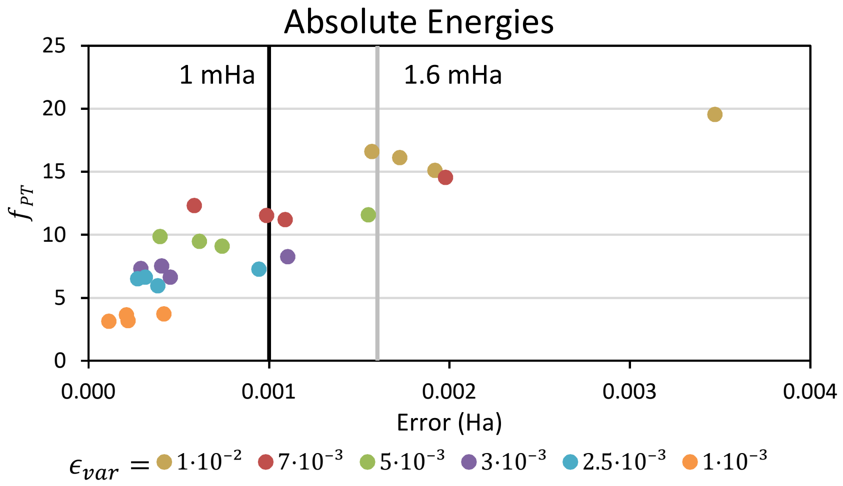


Figure D.1 The relationship between f_{PT} and absolute error for both absolute and relative energies in methylene.

D.2 Ethylene

D.2.1 ANO-L-pVDZ Data

For ethylene/ANO-L-pVDZ, steps of $0.1 \cdot 10^{-3}$ Ha were used from $\epsilon_{var} = 1 \cdot 10^{-2}$ to $1 \cdot 10^{-4}$ Ha to allow for accurate estimates of f_{PT}'s necessary for convergence. Figure D.2 shows f_{PT} behavior as ϵ_{var} decreases. As reported in Table 5.1, relative energies are converged to chemical accuracy with an average f_{PT} = 17.1% and to sub-mHa with an average f_{PT} = 10.4%. For absolute energies, chemical accuracy is obtained with an average f_{PT} = 12.9% and sub-mHa errors with an average f_{PT} = 10.3%.

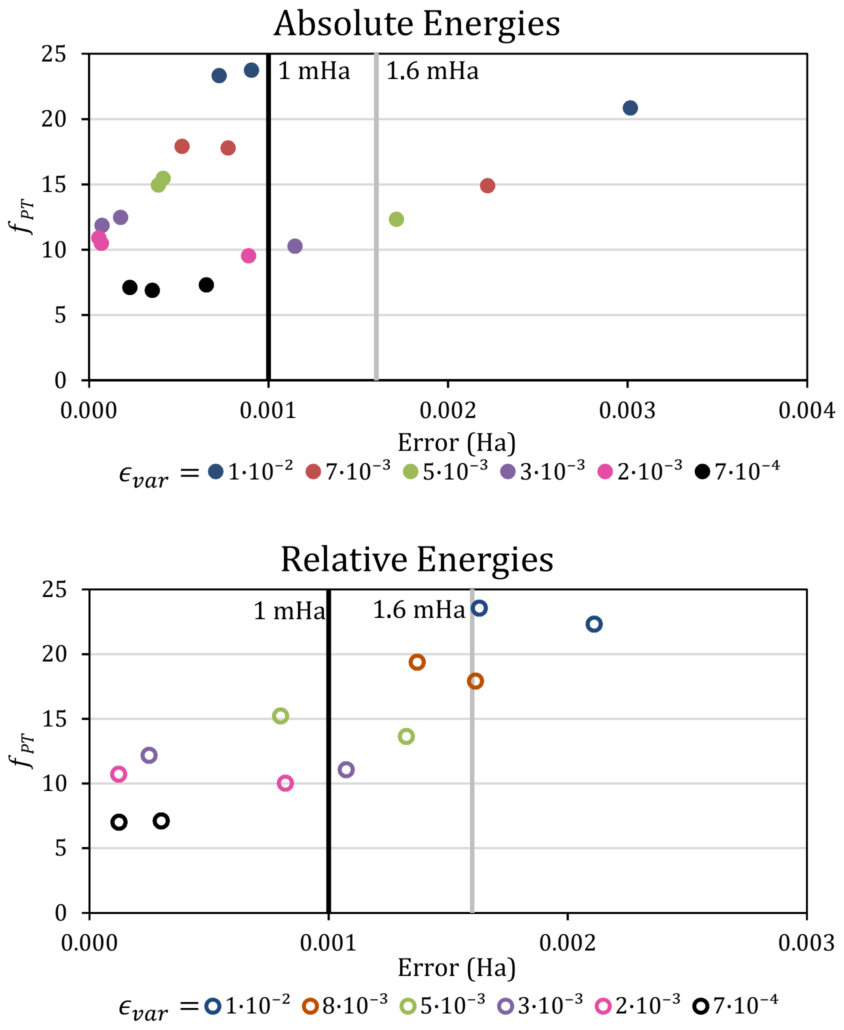


Figure D.2 The relationship between f_{PT} and absolute error for both absolute and relative energies in ethylene/ANO-L-pVDZ.

Table D.1 shows that ethylene/ANO-L-pVDZ energies agree well with past high-level calculations. The absolute energies are within a mHa of FCIQMC calculations with the same basis, and relative energies agree to within a mHa for FCIQMC, and is close to those computed with iFCI/6-31G*. CR-EOMCC(2,3)D results again show that more than triples excitations need to be accounted for coupled-cluster to accurately correlate 12 electrons, as was the case for ethylene/ANO-L-pVTZ.

Table D.1 Comparison of SHCI ethylene/ANO-L-pVDZ energies to literature values.

State	SHCI[a] (Ha)		FCIQMC (Ha)[b]		
1^1A_g	-78.3603(1)		-78.3599(1)		
1^1B_{1u}	-78.0536(1)		-78.0538(2)		
1^3B_{1u}	-78.1915(1)		-		
Gap	SHCI[a] (eV)	CR-EOMCC(2,3)D (eV)	FCIQMC (Ha)[b]	iFCI (eV)	Exp (eV)
1^1B_{1u}-1^1A_g	8.35	8.01	8.33 [b]	-	7.66[d]
1^3B_{1u}-1^1A_g	4.60	4.70	-	4.66 [c]	4.3-4.6[e]

[a] Perturbative energy at $\epsilon_{var} = 1 \cdot 10^{-5}$

[b] FCIQMC/ANO-L-pVDZ results from reference [1].

[c] iFCI/6-31G* results from reference [2].

[d] Reference [3].

[e] References [4–6].

D.2.2 ANO-L-pVTZ Extrapolations

Extrapolations are accomplished by fitting a line to the graph of the variational energies versus perturbative corrections. Table D.2 reports extrapolated energies obtained by fitting with different data sets ($\epsilon_{var} = [1 \cdot 10^{-3}, 1 \cdot 10^{-5}]$ and $[1 \cdot 10^{-4}, 1 \cdot 10^{-5}]$) and fitting equations (linear and quadratic). Specifically, we report the differences of these extrapolated energies with our most accurate SHCI calculation at $\epsilon_{var} = 1 \cdot 10^{-5}$. The extrapolated energies are extremely consistent, being at most 0.2 mHa away from the best SHCI value, indicating SHCI convergence.

Table D.2 Differences in ethylene/ANOL-pVTZ extrapolated energies with the best E_{SHCI} at $\epsilon_{var} = 1 \cdot 10^{-5}$. All values in Ha.

ϵ_{var} Range	Fitting Equation	1^1A_g	1^1B_u	1^3B_u
$1 \cdot 10^{-3}$ to $1 \cdot 10^{-5}$	Linear	-0.0001	-0.0001	-0.0002
	Quadratic	-0.0001	-0.0001	-0.0001
$1 \cdot 10^{-4}$ to $1 \cdot 10^{-5}$	Linear	0.0000	0.0000	-0.0001
	Quadratic	0.0001	0.0002	0.0001
Best SHCI ($\epsilon_{var} = 1 \cdot 10^{-5}$)		-78.4381	-78.1424	-78.2693

D.3 Ozone

D.3.1 cc-pVDZ Extrapolations

Extrapolations are accomplished by fitting a line to the graph of the variational energies versus perturbative corrections. Table D.3 reports extrapolated energies obtained by fitting with different data sets ($\epsilon_{var} = [1 \cdot 10^{-3}, 7 \cdot 10^{-6}]$ and $[1 \cdot 10^{-4}, 7 \cdot 10^{-6}]$) and fitting equations (linear and quadratic). Specifically, we report the differences of these extrapolated energies with our most accurate SHCI calculation at $\epsilon_{var} = 7 \cdot 10^{-6}$. The extrapolated energies are very consistent, being at most 0.4 mHa away from the best SHCI value, indicating SHCI convergence.

Table D.3. Differences in ozone/cc-pVDZ extrapolated energies with the best E_{SHCI} at $\epsilon_{var} = 7 \cdot 10^{-6}$. All values in Ha.

ϵ_{var} Range	Fitting Equation	1^1A_1	2^1A_1
$1 \cdot 10^{-3}$ to $7 \cdot 10^{-6}$	Linear	-0.0003	-0.0002
	Quadratic	-0.0004	-0.0003
$1 \cdot 10^{-4}$ to $7 \cdot 10^{-6}$	Linear	-0.0002	-0.0002
	Quadratic	0.0000	0.0000
Best SHCI ($\epsilon_{var} = 7 \cdot 10^{-6}$)		-224.9162	-224.7654

D.3.2 cc-pVTZ Extrapolations

Ozone/cc-pVTZ is the first system for which benchmark energies cannot be obtained with current computational resources, with the most accurate SHCI calculation at $\epsilon_{var} = 3 \cdot 10^{-5}$ leading to E_{PT} values of 15.8-25.8 mHa. Furthermore, Table D.4 shows that extrapolated energies are significantly different from the SHCI values, differing by up to 2.6 mHa. This indicates that the SHCI energies are not converged to FCI limit.

Table D.4 Differences in ozone/cc-pVTZ extrapolated energies with the best E_{SHCI} at $\epsilon_{var} = 3 \cdot 10^{-5}$. All values in Ha.

Geometry	ϵ_{var} Range	Fitting Equation	1^1A_1	2^1A_1
OM	$1 \cdot 10^{-3}$ to $3 \cdot 10^{-5}$	Linear	-0.0009	-0.0009
		Quadratic	-0.0012	-0.0014
	$1 \cdot 10^{-4}$ to $3 \cdot 10^{-5}$	Linear	-0.0006	-0.0007
		Quadratic	-0.0002	0.0004
	Best SHCI ($\epsilon_{var} = 3 \cdot 10^{-5}$)		-225.1379	-224.9858
RM	$1 \cdot 10^{-3}$ to $3 \cdot 10^{-5}$	Linear	-0.0009	-0.0016
		Quadratic	-0.0005	-0.0013
	$1 \cdot 10^{-4}$ to $3 \cdot 10^{-5}$	Linear	-0.0005	-0.0014
		Quadratic	0.0006	-0.0004
	Best SHCI ($\epsilon_{var} = 3 \cdot 10^{-5}$)		-225.0901	-224.8638
TS	$1 \cdot 10^{-3}$ to $3 \cdot 10^{-5}$	Linear	-0.0018	-0.0016
		Quadratic	-0.0022	-0.0006
	$1 \cdot 10^{-4}$ to $3 \cdot 10^{-5}$	Linear	-0.0019	-0.0003
		Quadratic	0.0026	-0.0011
	Best SHCI ($\epsilon_{var} = 3 \cdot 10^{-5}$)		-225.0496	-225.0482

D.4 Butadiene

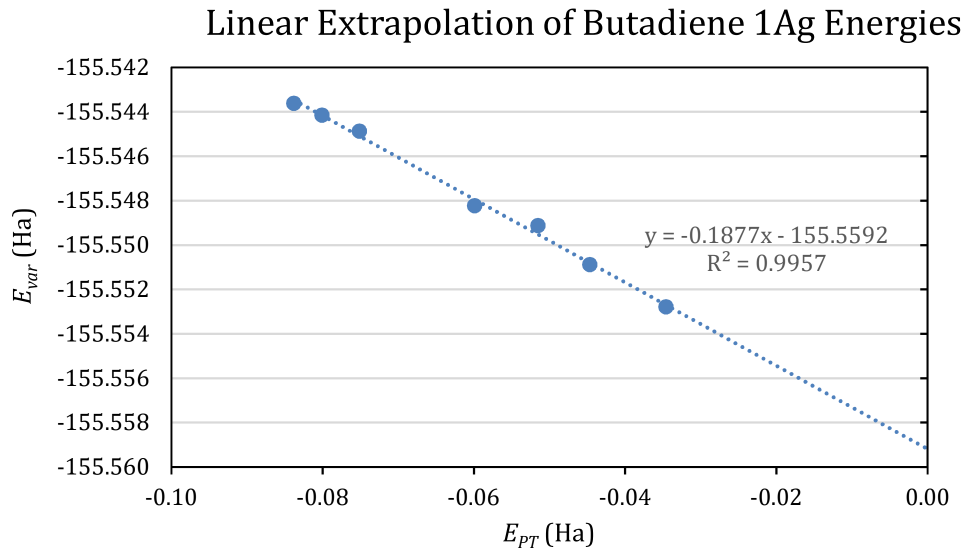


Figure D.3 Linear extrapolation of SHCI butadiene 1^1A_g energies.

D.5 References

1. Daday, C., Smart, S., Booth, G. H., Alavi, A. & Filippi, C. Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question. *J. Chem. Theory Comput.* **8,** 4441–4451 (2012).
2. Zimmerman, P. M. Singlet–Triplet Gaps through Incremental Full Configuration Interaction. *J. Phys. Chem. A* **121,** 4712–4720 (2017).
3. Mulliken, R. S. The excited states of ethylene. *J. Chem. Phys.* **66,** 2448–2451 (1977).
4. Moore, J. H. & Doering, J. P. Ion Impact Spectroscopy: Inelastic Scattering of 150–500-eV H^+ and H_2^+ from N_2, CO, C_2H_2, and C_2H_4. *J. Chem. Phys.* **52,** 1692–1699 (1970).
5. Moore, J. H. Investigation of the low energy singlet-triplet and singlet-singlet transitions in ethylene derivatives by ion impact. *J. Phys. Chem.* **76,** 1130–1133 (1972).
6. Van Veen, E. H. Low-energy electron-impact spectroscopy on ethylene. *Chem. Phys. Lett.* **41,** 540–543 (1976).
7. Sherrill, C. D., Leininger, M. L., Van Huis, T. J. & Schaefer, H. F. Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (TZ2P) basis. *J. Chem. Phys.* **108,** 1040 (1998).
8. Theis, D., Ivanic, J., Windus, T. L. & Ruedenberg, K. The transition from the open minimum to the ring minimum on the ground state and on the lowest excited state of like symmetry in ozone: A configuration interaction study. *J. Chem. Phys.* **144,** 104304 (2016).