Title

A State-Specific Complete Active Space Self-Consistent Field Approach for Strongly Correlated Electronic Excited States

Author

Hanscam, Rebecca

Publication Date

2023

Peer reviewed|Thesis/dissertation

A State-Specific Complete Active Space Self-Consistent Field Approach for Strongly Correlated Electronic Excited States

by

Rebecca Hanscam

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eric Neuscamman, Chair
Professor Martin Head-Gordon
Professor Joel Moore

Fall 2023

A State-Specific Complete Active Space Self-Consistent Field Approach for Strongly Correlated Electronic Excited States

Abstract

A State-Specific Complete Active Space Self-Consistent Field Approach for Strongly Correlated Electronic Excited States

by

Rebecca Hanscam

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Eric Neuscamman, Chair

Chemistry is teeming with light-driven processes such as photosynthesis and molecular switches, where electronically excited states are paramount to understanding the underlying mechanisms at play. Nevertheless, the appropriate treatment of effects such as strong electron correlation and orbital relaxations are beyond the simplifying assumptions of linear response theory and remains a challenge for theoretical modeling, consequently hampering the theory's ability to provide meaningful predictions and inform experiment. In this thesis, we present a novel approach with the flexibility and precision to overcome these obstacles.

To contextualize this work, we first discuss current excited state ansatzes and optimization methods, demonstrating their strengths and weaknesses in handling the challenges of modeling multi-reference systems like excited states. Using a generalized variational principle, we construct a fully excited-state-specific wave function optimization algorithm and demonstrate both its resilience to common optimization problems such as root flipping and its ability to locate and tightly converge excited state stationary points. Exhibiting both accuracy and reliability in calculating states with strongly correlated character and strong orbital relaxations, this approach is capable of providing state-specific insight in scenarios which elude existing methods, including some core, charge transfer and double excitations. We expand our study to explore the effect of these methodological improvements in comparison to a prominent traditional approach for geometry relaxations and potential energy surfaces for a photochemically relevant system, helping to elucidate the complexity of excited state energetics. The realization of this novel method constitutes significant steps forward for state-specific multi-reference approaches for the precise, accurate and robust modeling of electronically excited states.

Dedicated to the memory of Mike Hanscam

This one is for you, Dad.

Contents

List of Figures

List of Tables

Acknowledgments

Five years ago, I showed up to Berkeley barely knowing anything about theoretical chemistry; I was completely naive to its breadth and complex intricacies. Now, after many passes through Helgaker's "Big Purple Book", I largely feel the same way, but with an added appreciation for the true beauty of the origins of this field and its rapid expansion to fit the needs of so many different corners of chemistry. This thesis is my small contribution, it is a collective effort from those who came before me and everyone who supported my journey.

First and foremost, Professor Eric Neuscamman, who welcomed me to his group despite the steep learning curve I had yet to overcome. He trusted that I would still be worth the investment after a very tumultuous first year. He trusted that productivity for me sometimes meant working from a remote campsite or mountainside. And he trusted that I could dig myself out of the many tangles of optimization theory and matrix derivatives to contribute meaningfully to this field. A staunch advocate, a passionate teacher—willing to meet me wherever my understanding was at—and a persistent optimist, Eric embodied the exact mentorship I needed to succeed. This thesis did not come together smoothly, but through Eric's guidance, patience and expertise, the journey has made me a more confident scientist, person and mentor. His understanding was invaluable throughout the graduate school roller coaster ride, and during very difficult moments in my personal life. I am immensely grateful for all of his time and effort over the past five years. I would not be completing this degree without him.

I would like to thank my undergraduate mentors, Prof. Luc Boisvert, Prof. John Hanson and Prof. Carl Toews for teaching me how to be a scientist and feeding my curiosity for how chemistry and math intersect. Their mentorship and enthusiasm for teaching inspired me to continue down the path of academia. Although, I honestly do not miss working in a glovebox or a fume hood (sorry Luc and John!). Thank you to Professor Vale Molinero and Dr. Laura Lupi for introducing me to the power of supercomputers and sparking the beginning of my journey in the computational sciences. And finally, thank you to Professor Robert Szilagyi who has been both an advisor and a close friend over the years. His passion and drive for science is contagious, and the research opportunities he provided have opened many doors I didn't even know existed. I thank him for encouraging me to reach toward goals higher than I had aimed for, and believing in me even when I couldn't. He has been a consistent force in my life and I am grateful he urged me to apply to Berkeley, because (as usual!) he was right.

I wish to thank my qualifying exam committee: Prof. Martin Head-Gordon, Prof. Birgitta Whaley, Prof. Naomi Ginsberg, and Prof. Joel Moore. Despite the exam occurring at the very beginning of a pandemic and amid the first shelter-in-place orders, they were able to challenge me to think critically about my work in new ways. I thank them for their support and patience and Martin and Joel for taking the time to read this dissertation as part of my thesis committee.

I want to thank everyone in the Neuscamman Group for creating such an inclusive and supportive environment for this research to take place. Past group members Dr. Jacki Shea,

Dr. Leon Otis, Dr. Luning (Chris) Zhao, Dr. Beatrice Van der Goetz and Dr. Lan Tran for patiently teaching me the basics of quantum chemistry and walking me through the intricacies of the many electronic structure methods. My first year of graduate school was quite the struggle and I am grateful I had mentors to guide me through and help me start my research. My fellow cohort mates, including Dr. Scott Garner for always asking the difficult questions and being a reliable source of laughter, especially during our late-night library study sessions. I thank him for being my go-to source for symmetry related questions–I promise I will eventually relearn point groups! I thank Connie Robinson, Dr. Tarini Hardikar and Dr. Rachel Clune for their steadfast friendship and encouragement through both the toughest of times and the biggest of triumphs. Special thanks to Rachel for reminding me to eat, sleep and do laundry in addition to being an excellent sounding board when I needed to pivot from thinking about math to thinking about chemistry. My officemate Harrison Tuckman, who put up with far too many dog stories, I thank for always lending me his ear when I wanted to complain about whatever fresh source of frustration the day had brought forth, and laugh with me when the bug in my code was inevitably a simple addition error and rarely the complex math we covered our white boards with. I hope our paths cross again in the future, hopefully in an office without vines growing through the windows. A huge thanks to Sonja Tomasi and Trine Quady whose passion and tenacity for science inspire me every day and make me a better scientist. For all the camaraderie, debugging sessions, afternoon coffee walks, doggy play dates, late nights in the Gilman basement, Zoom calls throughout the pandemic and general shenanigans, thank you. I am proud of the tight-knit community the Neuscamman Group has become and I am deeply appreciative for the unwavering stability and support during the most challenging times. Their company over these last five years has been invaluable to me.

I also wish to thank members of the Martin Head-Gordon group: Diptarka Hait for many helpful conversations throughout the years, and Leonardo dos Anjos Cunha for being a consistent voice of reason when I was caught in the snarls of matrix derivatives.

I must also thank my very opinionated husky Kona for being a constant source of comic relief and for emphatically reminding me that the answer to every struggle is to take him for a walk (obviously).

I feel very lucky to have such an incredible support system, and that will always include my family. My sister, the first Dr. Hanscam, who is always quick to call me a science nerd and also help me navigate the frustrations of academia. My Mom, who has superhuman strength and compassion, and who is always available to listen when I need it. And finally, my Dad. Words are insufficient to capture how deeply I wish that he could be here to see me become the second Dr. Hanscam. He taught me to be curious about my environment, to explore the unknown, to ask questions and to never doubt the impact I could have on the world. He was a huge source of inspiration and encouragement throughout this academic journey and while these last few months without him have been painful, he will always be a part of my success. He would be the first to memorize the title of this thesis, to dissect each line and print off a copy to show anyone who would listen. Dad, you finally have your Pair-of-Docs!

Chapter 1

Introduction

1.1 Motivation

When a molecule is exposed to light it enters a higher energy state or excited state. For example when light energizes the molecules of plant leaf, they become more reactive than in the ground state or lowest energy state when in the dark. The reaction pathway of a photoexcited molecule often includes access to different routes and alternative products, involving crossings between potential energy surfaces that facilitate rapid conversion from reactant to product. Chemistry is full of such light-driven processes like photosynthesis and molecular switches, where geometric relaxations of such excited states play a critical role in determining the reaction pathway of a photoexcited molecule.[1, 2] Features of this potential energy surface (PES) such as the locations, depths and barrier heights between excited-state minima determine how a molecule responds and transforms when exposed to light. Other systems with similarly complex excited state character include carotenoids[3] and transition metal oxides.[4] Accurately modeling these processes to provide a meaningful prediction and inform experimental results requires an approach that can handle both the chemical and mathematical challenges posed by these systems.[5]

In modeling electronically excited states, we retain all the complications from modeling ground states but add a whole host of other challenges. When modeling a higher-energy system, it can be difficult to distinguish between local single excitations, multi-electron excitations, and charge-transfer states where the electron density shifts significantly from one region of a molecule to another.[6, 7] In many methods, one must resolve all lower-lying states before describing the excited state of interest is even possible, and choices made to accommodate descriptions of those states may be detrimental to describing the state in question.[8, 9, 10] Beyond the excited molecule itself, the molecular environment can play a large role in excited states' energies and properties, often necessitating the explicit inclusion of nearby protein structure or solvent.[11, 12, 13, 14, 15] Despite these challenges, progress is possible and ongoing thanks both to improved excited state optimization strategies[16, 17, 18, 19, 20] and better excited state wave function forms.[21, 22, 23, 24, 25, 26, 27] In this thesis, we present a theoretical method that distinguishes among a sea of energetic

eigenstates to accurately and reliably describe the orbital character of strongly correlated excited states: excited-state-specific complete active space self-consistent field theory using a generalized variational principle.

1.2 Schrödinger Equation

The time dependent Schrödinger equation is a fundamental equation governing quantum chemistry that determines how a wave function propagates in time and whose mathematical properties describe various physical properties of the system $|\psi\rangle$.

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle \tag{1.1}$$

In modeling excited states of molecules without heavy elements, it is often a good approximation to ignore relativistic effects[28] and to focus on the stationary states (those that do not change in time), which are given by the eigenvectors of the Hamiltonian operator. These can be evaluated by solving the time independent Schrödinger equation

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{1.2}$$

where $\hat{H}$ is the time independent Hamiltonian operator governing the physics of the system. For systems composed of atoms and electrons such as molecules, the non-relativistic time independent Hamiltonian operator $\hat{H}$ takes the form

$$\hat{H} = -\sum_{i}^{n}\frac{1}{2}\nabla_i^2 - \sum_{A}^{M}\frac{1}{2m_A}\nabla_A^2 - \sum_{i}^{n}\sum_{A}^{M}\frac{Z_A}{\left|\vec{r}_i - \vec{R}_A\right|} + \sum_{i>j}^{n}\frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{A>B}^{M}\frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|} \tag{1.3}$$

where n and M are the number of electrons and nuclei in the system, the lower case indices i, j track electrons and upper case indices A, B track nuclei, m is mass and Z is charge and finally $\vec{r}$ and $\vec{R}$ correspond to the positions of electrons and nuclei respectively. The five terms in $\hat{H}$ describe the kinetic energies of the electrons and nuclei as well as the potential energies of electron-nuclear, electron-electron, and nuclear-nuclear interactions.

1.3 Born-Oppenheimer Approximation

Further simplification of the Schrödinger equation comes with the Born-Oppenheimer Approximation. The mass of the nuclei are several orders of magnitude larger than the mass of the electrons, meaning that the timescale for motion of the nuclei is significantly slower than for electrons allowing the electrons to rapidly adapt to any change in nuclear geometry.[29] The Born-Oppenheimer Approximation recognizes that the nuclei can therefore be taken as stationary with respect to the electrons. Under this approximation, the Hamiltonian

operator reduces to

$$\hat{H}_{elec} = -\sum_{i}^{n} \frac{1}{2}\nabla_i^2 - \sum_{i}^{n}\sum_{A}^{M} \frac{Z_A}{\left|\vec{r}_i - \vec{R}_A\right|} + \sum_{i>j}^{n} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{1.4}$$

depending explicitly on the position of the electrons for a given nuclear geometry. The nuclear degrees of freedom are instead calculated for a given electronic PES using an effective potential $E_{elect}(\vec{R})$ for the electronic states on which the nuclei can move.

$$\hat{H}_{nuc} = -\sum_{A}^{M} \frac{1}{2m_A}\nabla_A^2 + E_{elec}(\vec{R}) \tag{1.5}$$

The nuclear and electronic contributions to the wave function are effectively separated, where the total wave function can be expressed as $\Psi(\vec{r},\vec{R}) = \Phi(\vec{R})\psi(\vec{r},\vec{R})$ such that the energy eigenvalues become sums over both contributions $E = E_{elec} + E_{nuc}$.

1.4 Diabatic and Adiabatic States

The Born-Oppenheimer potential energy surfaces on which the nuclei move are often referred to as adiabatic states,[30, 31] meaning those that diagonalize the electronic Hamiltonian. These states are not always chemically intuitive, however, as they can change character substantially as the nuclei move, for example between neutral and ionic states.[31, 32, 33] Both for intuition and in some quantum dynamics algorithms,[32, 34, 35, 36] it can instead be useful to construct electronic states that maintain their chemical character in the face of molecular motion. Such diabatic states can be formed by diagonalizing other operators, such as the dipole moment, in the basis of the adiabats[31, 30] and are sometimes found as the solutions to state-specific excited state methods.[16, 33]

As an example of adiabatic and diabatic states in a simple bond dissociation, consider the case of LiH, [32, 33, 37] shown in Figure 1.1. The adiabatic surfaces are the Hamiltonian eigenstates and, as such, show us how the excitation energy will change as a function of the bond stretch. However, neither adiabat has a simple chemical character. Near the equilibrium bond length, the lower adiabat is ionic and the upper adiabat is neutral (as seen for example by their Mulliken populations), but these characters switch upon dissociation, where the lower adiabat is neutral and the upper one is ionic. The two diabats, on the other hand, retain their ionic and neutral character, respectively, across the whole bond dissociation coordinate.

More mathematically, in such a two-state system spanning a two-dimensional state space, the wave function can be written as a linear combination of the diabatic states—in this case one is ionic and one is neutral

$$|\psi\rangle = c_I\,|\psi_I\rangle + c_N\,|\psi_N\rangle \tag{1.6}$$

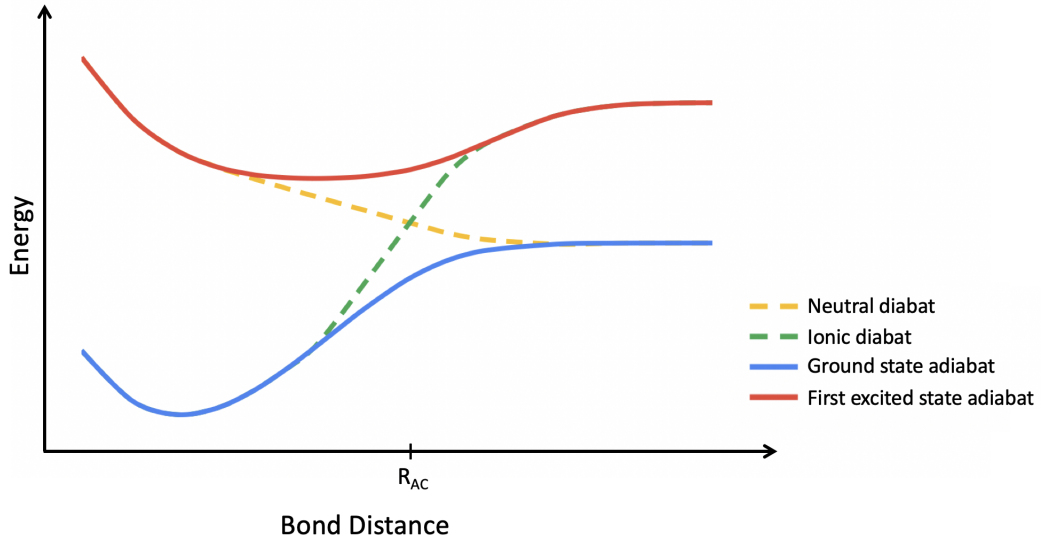


Figure 1.1. Simple representation of the LiH dissociation potential energy surface with an active space of 2 electrons in 2 orbitals, including the ground and first excited states, showing the adiabatic avoided crossing at R_{AC} and the ionic versus neutral diabats.

where $c_I \left|\psi_I\right\rangle$ is the contribution from the ionic diabat and $c_N \left|\psi_N\right\rangle$ the neutral diabat. Solving $\hat{H} \left|\psi\right\rangle = E \left|\psi\right\rangle$ where

$$\hat{H} = \begin{bmatrix} \langle\psi_I|H|\psi_I\rangle & \langle\psi_I|H|\psi_N\rangle \\ \langle\psi_N|H|\psi_I\rangle & \langle\psi_N|H|\psi_N\rangle \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \tag{1.7}$$

gives the energies on the adiabatic potential energy curves, accounting for the coupling between the two diabats due to the non-zero off-diagonal elements in $\hat{H}$

$$E_{\pm} = \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}\sqrt{(H_{11} - H_{22})^2 + 4|H_{12}|^2}. \tag{1.8}$$

For these two curves to cross, or in other words for the resulting energies to be equal, it must be true that $H_{11} = H_{22}$ and $H_{12} = 0$. With only one geometric degree of freedom (the Li-H bond length), both conditions cannot be met and these states instead display an avoided crossing at R_{AC}. In poly-atomic systems with additional degrees of freedom, this no-crossing rule no longer applies and the Born-Oppenheimer surfaces are free to cross in what are referred to as conical intersections.[30, 31, 38]

1.5 Second Quantized Electronic Hamiltonian

Up this point the Hamiltonian has been defined in real space, however it is often much more convenient to project it into the large-but-finite vector space occupation numbers known as

Fock space.[29] To do so, we define Slater determinants that each contain N orthonormal molecular orbitals ϕ, each of which is a linear combination of one-electron basis functions χ_p.

$$|\phi_1\phi_2\dots\phi_N\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \cdots & \phi_N(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \cdots & \phi_N(\vec{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\vec{r}_N) & \phi_2(\vec{r}_N) & \cdots & \phi_N(\vec{r}_N) \end{vmatrix} \tag{1.9}$$

Here $\vec{r}_i$ represents the ith electron's combined spatial and spin coordinate, and the constant out front ensures the overall determinant is normalized. The different Slater determinants can be relabeled using occupation number vectors,

$$|\vec{n}\rangle = |n_1 n_2 \dots n_N\rangle \tag{1.10}$$

in which n_p is 1 if orbital ϕ_p is occupied (present in the determinant) and 0 if unoccupied (not present). If the set of orbitals $\{\phi\}$ from which it is built are orthonormal, this many-electron basis for Fock space will be as well.[29]

In the occupation number basis we can write the Hamiltonian as[29]

$$\hat{H} = \sum_{pq}^{N_{basis}} \sum_{\sigma\in\alpha,\beta} h_{pq}\hat{a}^\dagger_{p\sigma}\hat{a}_{q\sigma} + \frac{1}{2}\sum_{pqrs}^{N_{basis}} \sum_{\sigma,\tau\in\alpha,\beta} (pq|rs)\hat{a}^\dagger_{p\sigma}\hat{a}^\dagger_{r\tau}\hat{a}_{s\tau}\hat{a}_{q\sigma} + h_{nuc} \tag{1.11}$$

where $\hat{a}^\dagger_{p\sigma}$ is a creation operator that places a σ spin electron in orbital ϕ_p and $\hat{a}_{q\tau}$ is an annihilation operator that removes a τ spin electron from orbital ϕ_q. The one-electron integral contains two pieces, the electron kinetic energy operator and the electron-nuclear Coulomb operator.

$$h_{pq} = -\frac{1}{2}\int \phi^*_p(\vec{r}_1)\nabla^2\phi_q(\vec{r}_1)d\vec{r}_1 - \int \phi^*_p(\vec{r}_1)\sum_A \frac{Z_A}{\left|\vec{r}_1 - \vec{R}_A\right|}\phi_q(\vec{r}_1)d\vec{r}_1 \tag{1.12}$$

The two-electron integrals

$$(pq|rs) = \iint \phi^*_p(\vec{r}_1)\phi_q(\vec{r}_1)\frac{1}{|\vec{r}_1 - \vec{r}_2|}\phi^*_r(\vec{r}_2)\phi_s(\vec{r}_2)d\vec{r}_1 d\vec{r}_2 \tag{1.13}$$

define the electron-electron repulsion operator and, lastly, h_{nuc} is the nuclear-nuclear repulsion energy.

1.6 Overview of Electronic Structure Methods

Even with the time-independent Schrödinger equation and Born-Oppenheimer Approximation in place, the electronic problem remains too complex to be solved analytically for multi-electron systems in either the ground or an excited state. Further simplifications must be

made and many methods exist that use various types and levels of approximations. At a high level, methods can be divided between the categories of density functional based[39] and wave function based[28, 29] approaches. Density Functional Theory (DFT) formulates the energy as a functional of the electron density and is the most widely used of the two categories of the methods. While it has the significant advantage of a relatively low computational cost, the accuracy of DFT is sensitive to the choice of approximate density functional.[40] Improving the quality of the approximate functionals is a large focus in the field, but doing so systematically has proven difficult and continues to motivate research into wave function methods.

1.6.1 Hartree-Fock Theory

In ground state quantum chemistry, there exist multiple systematically improvable wave function hierarchies that start from a single determinant, such as configuration interaction theory[41] and coupled cluster theory,[42] but to work well these typically must start from a high quality single-determinant description of the system.[29] Typically, this starting point comes from Hartree-Fock (HF) theory,[28] which minimizes the determinant's energy by optimizing its molecular orbital shapes. The HF ansatz can be written as

$$|\Psi_{HF}\rangle = e^{-\hat{X}} |\Phi\rangle \tag{1.14}$$

where $|\Phi\rangle$ is a single Slater determinant formed from orthonormal molecular orbitals $\{\phi_p\}$ typically expressed as linear combinations of atomic orbitals $\{\chi_i\}$.

$$\phi_p(\vec{r}) = \sum_{\mu} c_{\mu p} \chi_\mu(\vec{r}) \tag{1.15}$$

The unitary orbital rotation operator $e^{-\hat{X}}$ is defined by an anti-Hermitian matrix $\boldsymbol{X}$.

$$\hat{X} = \sum_{pq} X_{pq} \hat{a}_p^\dagger \hat{a}_q \tag{1.16}$$

The ground state variational principle (GSVP)—discussed further in Section 1.7.1—states that the expectation value of the Hamiltonian using any approximate wave function must be greater or equal to the lowest Hamiltonian eigenvalue. While one could minimize the energy of the wave function in Eq. (1.14) directly, alternatively one can construct a Lagrangian minimization to find the energy via the GSVP by changing the wave function parameters $c_{\mu p}$ under the constraint that the final orbitals are orthonormal.

$$L = \frac{\langle \Psi_{HF} | \hat{H} | \Psi_{HF} \rangle}{\langle \Psi_{HF} | \Psi_{HF} \rangle} - \sum_{pq} \epsilon_{pq} \left(\langle \phi_p | \phi_q \rangle - \delta_{pq} \right) \tag{1.17}$$

The set of Lagrange multipliers $\{\epsilon\}$ become the HF molecular orbital energies. Finding a stationary point where all $\frac{\partial L}{\partial c_{\mu p}} = 0$ is equivalent to the generalized eigenvalue problem given

by the Roothaan-Hall equations[43, 44, 28]

$$\boldsymbol{FC} = \boldsymbol{SC}\epsilon \tag{1.18}$$

where $\boldsymbol{C}$ is the LCAO coefficient matrix that converts the atomic basis functions into molecular orbitals, $\boldsymbol{S}$ is the atomic orbital overlap matrix, and $\boldsymbol{F}$ is the one-electron mean-field Hamiltonian known as the Fock matrix that depends on the orbital shapes. This generalized eigenvalue problem is solved iteratively by updating the molecular orbitals in what is referred to as the self-consistent field (SCF) method. Generally speaking, the SCF method involves solving the Roothaan-Hall equation to obtain a more accurate orbital representation and then using these orbitals to set up a new Fock operator, repeating this process until the relative change is negligible.[28]

For most systems, this method of calculating a ground state representation typically recovers about 99% of the total electronic energy. The difference between the true electronic energy of a system and the HF energy—that remaining 1%—is defined as electron correlation energy and arises from effects beyond the statistically independent nature of the mean-field approach.[45] That remaining 1% of the energy is crucial for accurately calculating many molecular properties and can cause as large as 10's of kcals/mol errors in reaction energies.[46] Correlation energy is typically discussed in two different forms, strong and weak correlation. Weak correlation is small corrections to the energy resulting from the correlation of electron movement within the qualitative picture provided by HF. More advanced theories higher up in the hierarchy of methods aim to recover enough of the missing correlation energy from a HF starting point to become quantitatively accurate. Strongly correlated systems are instead described by multiple electronic configurations, as in those systems the single determinant approach of HF is qualitatively incorrect. Strong correlation results in large changes to the HF electronic structure and additional electronic configurations aside from HF are major contributions to the wave function. Such systems are referred to as having large multi-reference character.

1.6.2 Configuration Interaction

A straightforward way of improving on HF and recovering correlation effects is moving to a multi-determinant ansatz that considers every possible excitation from the Aufbau configuration. Starting from the HF molecular orbitals and including all possible electronic configurations forms a complete basis, such that an exact wave function for a system is the Configuration Interaction (CI) wave function

$$|\Psi_{CI}\rangle = c_0 |\phi_0\rangle + \sum_{ia} c_i^a \hat{a}_a^\dagger \hat{a}_i |\phi_0\rangle + \sum_{ijab} c_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i |\phi_0\rangle + ... \tag{1.19}$$

Here, orbitals that are occupied in the Aufbau configuration are indexed by i and j and the unoccupied orbitals by a and b. The determinant coefficients c for each respective electronic configuration $|\phi\rangle$ form a weighted sum of all possible single, double, triple, etc

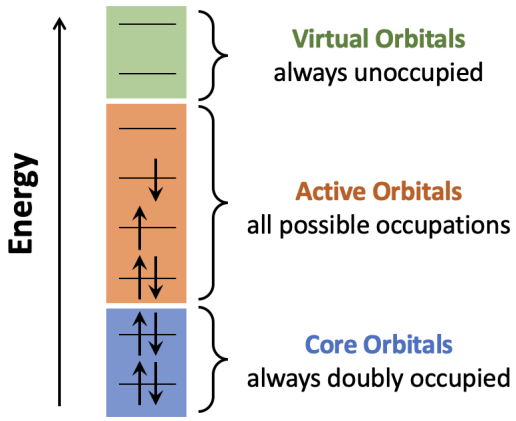


Figure 1.2. Schematic of CASSCF orbital spaces.

excitations off the Aufbau configuration. For small systems, the full expansion of this sum can be used to calculate an energy exact up to the limit of the basis set. This is a method referred to as Full CI (FCI). However FCI scales exponentially with system size and quickly becomes prohibitively computationally expensive beyond the smallest molecules and simplest basis sets.[29] In practice, the FCI wave function is often truncated at a particular level of excitation, such as CIS or CISD where only single excitations are included or single and double excitations respectively. While these truncated CI methods retain only the most important components of the wave function at only a polynomial cost, anything beyond CIS loses size consistency and size extensivity causing degrading accuracy as the system size increases.[47]

1.6.3 Complete Active Space Self-Consistent Field

A different approach to approximating the FCI wave function is applying FCI on only a subset of the orbitals called an active space, creating a complete active space (CAS) wave function.[48, 49, 50, 51, 29] The CAS wave function partitions the orbitals into three subspaces: core orbitals, active orbitals, and virtual orbitals. As shown in Figure 1.2, the core orbitals are doubly occupied low energy orbitals that are not expected to contribute significantly to the excited states of the system. In contrast, the active orbitals are those most likely to participate in the excitations of the system and will typically include both the highest energy occupied molecular orbitals and the lowest energy unoccupied molecular orbital. The CAS wave function is built from a linear combination of the reference determinant and all possible excitations within the set of active orbitals. This leaves the high-energy virtual orbitals that remain unoccupied from the reference. While active spaces only include the most chemically important orbitals, results of a CAS calculations are highly sensitive to the choice of orbitals included in this space.[52, 53] It seems natural to assume the more active

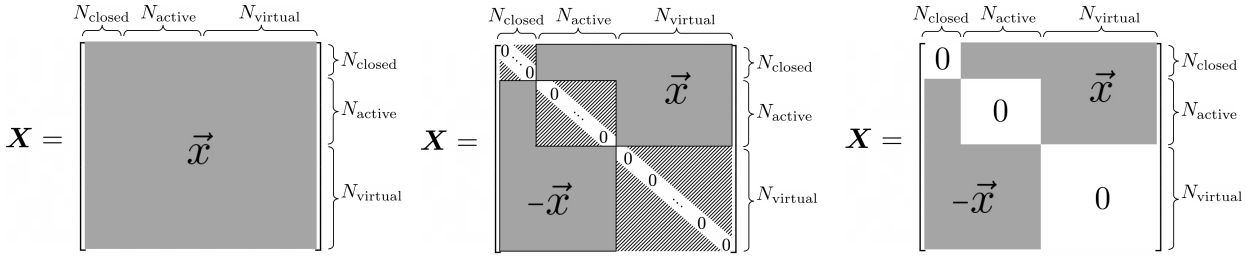


Figure 1.3. The orbital rotation coefficient matrix where the solid shaded regions represent nonzero variational parameters and the striped triangles identify the energy-invariant parameters. The block structure of $\boldsymbol{X}$ with no assumptions made is shown on the left, when anti-Hermiticity is assumed in the middle, and only non-redundant parameters on the right.

orbitals the better, but traditional FCI algorithms and conventional CASSCF approaches become intractable at a size of about 20 orbitals with 20 electrons.[54] While recent developments for methods like selected CI,[27] density matrix renormalization group,[24, 55] and full configuration interaction quantum Monte Carlo[20] are pushing the active space limit closer and closer to over a hundred orbitals,[56] chemical intuition and understanding remain highly valuable in constructing an effective active space and an accurate CASSCF treatment.

Unlike FCI and the truncated CI methods, CAS self-consistent field (CASSCF) improves its description via orbital relaxations, in particular the core-active, core-virtual, and active-virtual orbital rotations that affect the energy. Thus, CASSCF optimization methods must contend with both the excitation weights and the molecular orbital parameters. [48, 49, 50, 51, 29] Similar to HF, the CASSCF wave function can be written as

$$|\Psi_{CAS}\rangle = e^{\hat{X}} \sum_I c_I |\phi_I\rangle \tag{1.20}$$

where $|\phi_I\rangle$ represents a Slater determinant and c_I the corresponding CI coefficient. The orbital rotation matrix $\boldsymbol{X}$ described in Eq. (1.16) is anti-Hermitian and thus defined by its upper triangle, as depicted in Figure 1.3.[29] Additionally, rotations between orbitals within a given orbital subspace are energy invariant and already accounted for by the CI expansion, and therefore are not included in the orbital parameterization to avoid redundancy.

Because of the orbital optimization, the resulting CAS wave function is considerably more flexible than traditional truncated CI methods, particularly for strongly correlated systems with degenerate or nearly degenerate electronic configurations. CASSCF is often used as a qualitatively accurate reference wave function for higher-level multi-reference methods to recover more of the weak correlation missed by CASSCF from the orbitals outside the active space. One approach for addressing this is to perform perturbation theory using the CASSCF wave function as a reference, leading to the CASPT2 method[57, 58] which is discussed further in Section 1.8.4.

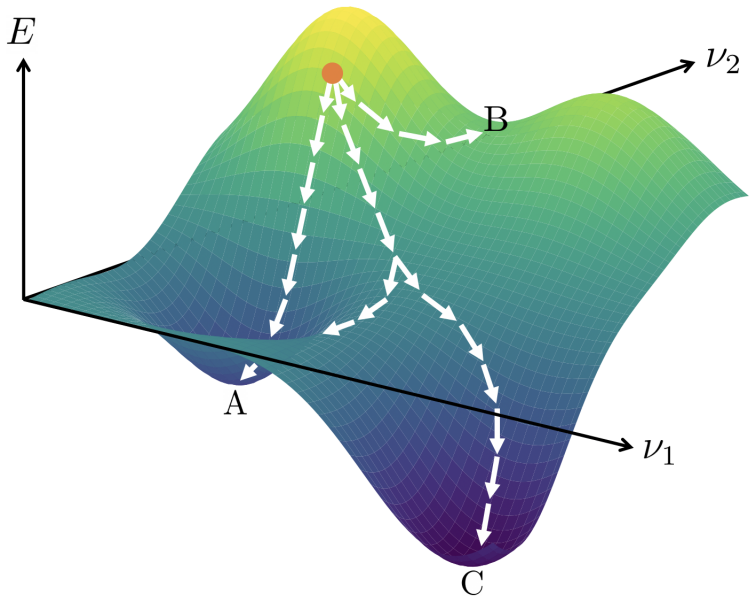


Figure 1.4. Idealized representation of the PES between the orbital shapes and the electronic configuration. Possible optimization pathways are shown, where routes to stationary states A and C would be considered variational collapse when targeting state B.

1.7 Optimization Strategies

With an approximate wave function ansatz in hand, its parameters must then be optimized to resemble the exact wave function as closely as possible in order to converge to an approximate Hamiltonian eigenstate. Excited states in particular are challenging for optimization methods as they are saddle points in the FCI space,[59, 60, 61] which are more difficult to characterize than minima. As saddle points, excited states are inherently more complicated for numerical optimization techniques and are more prone to optimization issues like variational collapse, discussed further in the following section.

1.7.1 Variational Principles

One approach to optimizing the wave function parameters is to minimize the energy functional

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0 \tag{1.21}$$

where the energy expression built from the current wave function $|\Psi\rangle$ is lower bounded by the exact ground state energy E_0. Because the wave function is a linear combination of Hamiltonian eigenstates and the exact ground state energy of the system is strictly less than or equal to the remaining eigenvalues of $\hat{H}\,|\phi_i\rangle = E_i\,|\phi_i\rangle$, it can be shown that the

current energy of $|\Psi\rangle$ must be greater or equal to E_0. [62] Optimizing the parameters of an approximate wave function to the global minimum of its energy function will locate the ground state energy and wave function for that approximate ansatz. This is what is known as the ground state variational principle, as mentioned previously in Section 1.6.1.

The GSVP is widely used in the optimization of approximate wave function ansatzes in the hierarchy of methods designed to calculate ground states, such as HF, CISD, and density matrix renormalization group.[63, 29, 28] Variational collapse however, prevents the straight forward use of the GSVP to locate higher energy states. Variational collapse occurs during an optimization when the wave function significantly changes character from the initial state and optimizes to an altogether different energy stationary state.[64, 7] Consider the energy surface in Figure 1.4 depicting an energy minimization of a wave function dependent on the variational parameters ν_1 and ν_2. Even if the initial wave function, shown by the orange circle, is relatively close in parameter space to the targeted excited state (state B) that stationary point is a saddle point and is not a minimum of the energy functional, much less the global minimum. There are numerous more probable routes the optimization could take that instead lead to states A or C. When targeting an excited state, variational collapse is a realistic hurdle when minimizing the GSVP without other considerations for excited states.

Single determinant ansatzes like Δ-SCF are built atop the GSVP with added flexibility to handle excited states in a ground state HF or DFT optimization.[65] In this approach, electrons are placed in non-Aufbau configurations by swapping in the virtual orbitals involved in the desired excitation. However it is not always clear where to place electrons when the energy ordering of the orbitals change or the orbital shapes shift significantly upon relaxation, where differing decisions can lead to disparate results for the final outcome of the calculation. In addition, this approach can have limited flexibility for systems with more than one configuration state function. In an effort to aid the frustrating optimization issues demonstrated by Δ-SCF, a method of systematically tracking the non-Aufbau electronic configurations through the SCF procedure was introduced, referred to as the maximum overlap method (MOM).[66, 67] In this approach the orbitals fresh out of an SCF cycle are projected onto the occupied orbitals from the previous iteration and the set of orbitals with the largest overlap are selected for the next SCF cycle. More than just a Δ-SCF method, MOM can be applied to many other SCF ansatzes as well, including coupled cluster[68] and CASSCF,[16] and used in direct optimization setups.[69] The pairing of MOM with Δ-SCF-DFT has successfully identified atomic Rydberg states,[67] intermolecular charge transfer excitations,[67] and small molecule core excitations.[70] While MOM reduces the frequency of optimization problems, it has also been shown to fail due to variational collapse and so may benefit from incorporating additional state-targeting methods.[66, 16]

When targeting an excited state, instead of attempting to find an energy stationary point, one can optimize the wave function by minimizing an alternative function whose global minimum corresponds to the desired excited state. One such excited state variational

principle (ESVP) takes the form

$$W = \frac{\langle\Psi|(\omega - \hat{H})^2|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{1.22}$$

where ω is the approximate energy of the desired eigenstate.[71, 72] While the GSVP quantifies how low the energy is in an absolute sense, W quantifies a measure of the energetic "distance" between the current wave function and the targeted excited state. For an exact wave function, the global minimum of W occurs at the energy eigenstate with energy nearest to ω, regardless of whether that energy is above or below ω.[21]

Similar to variational principle minimization, an energy stationary point search can be performed by minimizing functions dependent on the gradient of the energy such that $\nabla_{\vec{\nu}} E = 0$ at convergence with respect to all variational parameters ν_i. Square gradient minimization (SGM) is one such approach that seeks to minimize the square of the gradient of a given variational principle with respect to the wave function parameters. [73]

$$\Delta_{SGM} = |\nabla E|^2 \tag{1.23}$$

As an extension of the geometric direct minimization method,[74] SGM uses a finite difference formulation that is more numerically efficient than previous approaches and is applicable to many approximate wave function ansatzes. Eq. (1.23) has local minima at each energy stationary state with respect to the current wave function, however for approximate ansatz there may exist additional minima including some that are not energy stationary points. [33] As all Hamiltonian eigenstates zero out the energy gradient, SGM must rely on the initial wave function guess starting the optimization within the basin of convergence of the desired excited state, but this issue can be mitigated by incorporating state-targeting properties in addition to the energy gradient within an expanded objective function.

The generalized variational principle (GVP) uses both properties and energy gradient minimization to define a more state-specific guide for optimizing excited states.[21] The GVP uses a set of property values to uniquely describe the targeted state, defining a vector of property deviations $\vec{d}$ from a set of operators and their desired expectation values. At the end of the optimization, this vector should have values close to zero. Typically included in $\vec{d}$ is an energy difference term $\langle\hat{H}\rangle - \omega$ that is motivated as an approximation to the ESVP[21, 16] and describes the difference between the current wave function energy and a guess for the energy of the targeted excited state. The GVP can be written as

$$L = \mu|\vec{d}|^2 + (1 - \mu)|\nabla E|^2 \tag{1.24}$$

where μ controls the weightings of each term throughout the optimization. Properties included in $\vec{d}$ help guide the optimization to the desired basin of convergence, after which μ is set to zero and $|\nabla E|^2$ ensures the resulting wave function represents an energy stationary point.

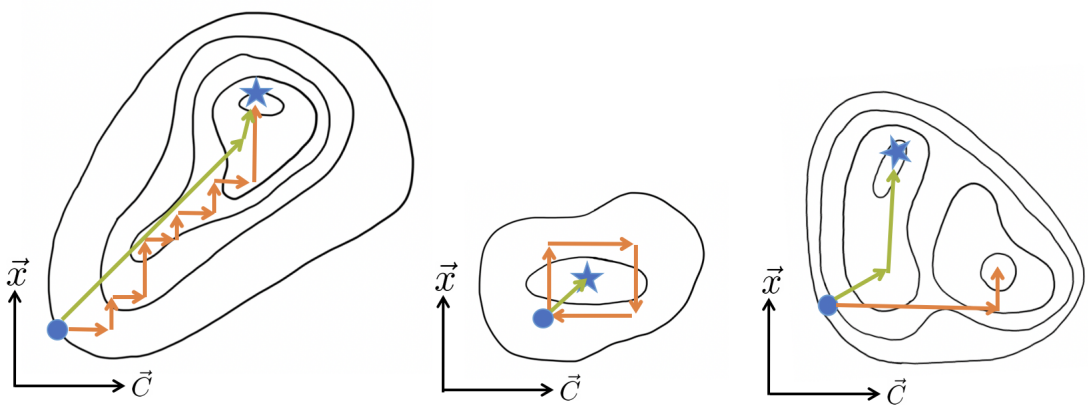


Figure 1.5. Schematic 2D contour plots of simple surfaces demonstrating two-step (orange) versus direct optimizations (green) in various convergence scenarios. Each optimization pathway begins at the blue circle with the objective being the blue star.

1.7.2 Direct Optimization versus Two-Step

In CASSCF, there are two components of the approximate wave function ansatz of concern: updating the CI coefficients and relaxing the molecular orbitals. Commonly these two problems are handled separately in order to take full advantage of efficient optimization strategies individualized to the unique characteristics of each parameter set.[75, 76, 52, 77, 53, 78, 79] Truncated CI methods and CASCI rely solely on the input orbital shapes and subsequently uses Krylov subspace methods[80] such as Davidson[81] to diagonalize the Hamiltonian. CASSCF on the other hand needs to relax the orbitals in addition to updating the CI.

One approach is to optimize the two parameter sets individually, where the orbital shapes are updated via an SCF procedure or an augmented Hessian approach,[17, 82, 83] then the CI Hamiltonian is rediagonalized in the new orbital basis and the process is repeated until convergence. This is referred to as a two-step approach, which tailors the optimization of each parameter set to take full advantage of their respective structure. However, with no passing of information between the parameter sets at any given step, the two-step process is often constrained to taking only small steps in parameter space, and even so can easily overstep and cause convergence issues.[17, 84] Alternatively a direct optimization approach updates both parameter sets simultaneously, hopefully making it possible to converge more efficiently, directly and accurately.[80] Several examples of this are shown in Figure 1.5 and include the classic limit cycle as well as the scenario where two-step oversteps into a different basin and cannot recover. Both are cases where a direct approach would be crucial to converging to the desired state.

Ground state CASSCF codes don't typically use a naive two-step optimization strategy and instead will include at least partial coupling between the orbital and CI coefficient pa-

rameter sets. For example, Werner-Knowles approaches to MCSCF optimization combines second-order quasi-Newton descent with CI diagonalization such that each update of the CI coefficients or the orbitals includes information on how the other parameter set will respond.[17, 84] Partial CI-orbital coupling and direct fully coupled approaches accelerate convergence rates and achieve convergence in cases where a two-step approach would not (Figure 1.5).[17, 84, 85] Each of these optimization procedures use variations on re-diagonalization of the Hamiltonian to update the CI coefficients, and consequently, successful convergence is dependent on the method's ability to track the relevant CI root throughout the optimization, even if the ordering of roots changes due to orbital relaxations.

1.8 Excited State Methods

There is a well-developed hierarchy of methods built atop Hartree Fock that recover progressively more electron correlation where each method was developed originally for the primary purpose of studying ground states of molecular systems. Excited states however, play a key role in unraveling the mechanisms of light-driven processes that often involve more complex wave functions and challenging features on their potential energy surfaces that ground state methodologies are unable to capture. The importance of excited states in this area motivates the development of accurate and effective theoretical models designed with excited-states in mind. Approaches to such methods fall broadly into two categories, linear response theories that aim to build excited state descriptions from a ground state reference and state-specific ansatzes that eschew ground state involvement. To put the opportunity afforded by the GVP for a state-specific approach into context, let us first discuss existing methodology for excited states, both state-specific and otherwise.

1.8.1 Linear Response Theories

Laboratory experiments often involve probing how a system responds when it is driven away from equilibrium. In similar manner, linear response theory utilizes small perturbations to move a molecule away from its ground state equilibrium. The wavefunction evolves to become a linear combination of multiple electronic states such that excited state properties can be measured. This approach is applicable in both wavefunction and density-based methods, however when combined with an approximate ansatz the implicit assumption that in this tangent space the excited states share much in common with the ground state often breaks down for states or mixtures of states with significant multiply excited character.[86]

Equation of motion coupled cluster (EOM-CC) theory is a popular post Hartree-Fock wave function approach to recovering weak correlation in excited state calculations and is a balanced combination of accuracy and computational affordability in many cases.[87, 88, 89, 90] The coupled cluster wave function[29]

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}} |\Psi_0\rangle \tag{1.25}$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots = \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i + \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j + \ldots \tag{1.26}$$

is built similarly to truncated CI methods, where higher-levels of cluster truncation can provide high accuracy single valence excitation energies in small molecules.[89] The doubles cluster operator $\hat{T}_2$ describes electron pair interactions while the singles cluster operator $\hat{T}_1$ relaxes the orbitals in response to the potential created by those interactions. To further perturb the wave function away from the ground state, the EOM-CC parameterization introduces a linear excitation operator

$$|\Psi_{\text{EOM-CC}}\rangle = \hat{R} e^{\hat{T}} |\Psi_0\rangle \tag{1.27}$$

$$\hat{R} = \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i + \frac{1}{4} \sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j + \ldots \tag{1.28}$$

that, as it commutes with $\hat{T}$, can be though of as either acting on the HF reference before the application of the cluster operator or as acting on the whole correlated CC ground state.

Several limitations arise from the EOM-CC setup. First since the cluster operator $\exp[\hat{T}]$ was optimized for the ground state, its electron correlation description for $\hat{R}|\Psi_0\rangle$ is less than optimal.[91] Second, the EOM-CC wave function, especially at the singles and doubles "CCSD" level, is highly dependent on the reference orbitals and wave function,[92] which can cause distinct and non-systematic errors leading to an unbalanced description of excited states.[93] For example, EOM-CCSD cannot achieve any orbital relaxation following a double excitation, as doing so even to first order would require triple excitations. Perturbative triples corrections offer some orbital relaxation in double excitations,[90] but they can still struggle to capture the full relaxation effects following charge transfer excitations.[93]

Time-dependent density functional theory (TD-DFT) is another widely used approach for modeling excited states. Building on ground state DFT, it has been highly successful in local single excitations and many, many absorption spectra studies.[94] Typically formulated as a linear response approach, TD-DFT analyzes the response of the time-independent ground state electron density to a time-dependent external electric field.[95] However the most common formulations are known to be inaccurate for double excitations and charge transfer states.[96, 97, 98, 99] The adiabatic approximation allows only single excitations to appear in the TD-DFT formalism, so higher-level excitations are entirely missing.[100, 101] TD-DFT chronically underestimates CT excitation energies and only reproduces the asymptotic $\frac{1}{R}$ behavior of particle-hole attraction when range-separated hybrid functionals are employed.[98, 99, 100] In addition, the same issues with predictive power in ground state DFT stemming from its sensitivity to functional choice carry over into the excited state method.[102, 103] Progress has certainly been made, for example via optimal tuning of range-separated functionals, [104] but doubly excited and CT states remain major challenges for TD-DFT.

1.8.2 Excited State Mean Field Theory

In parallel to what exists for ground states, some recent work has focused on developing a hierarchy of methods for state-specific excited state modeling, ranging from mean field[105] to low-level correlation[106, 107] to multi-reference theory.[19, 16] The base of the hierarchy is excited state mean field theory (ESMF). The ESMF ansatz is a CI wave function of single excitations

$$|\Psi_{\text{ESMF}}\rangle = e^{\hat{X}} \sum_{ia} c_i^a \hat{a}_a^\dagger \hat{a}_i \, |\phi_0\rangle \tag{1.29}$$

with state-specific orbital relaxation via minimization of a Lagrangian built using the GVP in Eq. (1.24).[105, 108, 109] It was this wave function ansatz that provided the playground in which the GVP was developed, showing its powerful scope for targeting excited states by energy, spin and Mullikan charge.[21] Analogous to the role of Hartree-Fock in the ground state, ESMF serves as a foundation for building state-specific versions post-HF theories such as excited-state-specific second-order perturbation theory (ESMP2)[106] and state-specific coupled cluster theory.[107] A benchmarking study of ESMP2 highlights the advantages of using the GVP-based state-specific approach, showing significant improvement over EOM-CC and a wide variety of TD-DFT approaches for singlet excitations in small π systems.[110] All of these methods are single-reference ansatz that recover weak correlation off a mean-field reference. A similar opportunity exists in more strongly correlated settings and is explored in this thesis, where in the same spirit as ESMF, a state-specific CASSCF method using the GVP could act as the foundation for a state-specific multi-reference hierarchy of methods.

1.8.3 Excited State CASSCF

Turning to multi-reference methods for the calculation of excited states, new optimization challenges arise, namely state-specific orbital optimization for a diversity of excitations and tracking the desired state through the optimization. As discussed in Section 1.6.3, there are two components to a CASSCF wave function optimization: description of the electronic configuration (CI vector) and orbital relaxation. When the Hamiltonian is diagonalized within the active space, we get an energetic hierarchy of states in which the ground state, first excited state, second excited state, etc are well defined. However, the relative energy ordering of these states is not necessarily maintained throughout the optimization. This problem is referred to as root flipping. When the target of a CASSCF calculation is an excited state and thus must contend with root flipping, there are several common approaches to optimizing the wave function. These range from simple diagonalization of the Hamiltonian to more complex state-tracking methods,[16, 67, 111] each using the two-step approach discussed in Section 1.7.2. To understand the challenges of existing two-step methods and to see why a one-step formulation may have advantages, let us consider a series of examples from the recent literature.

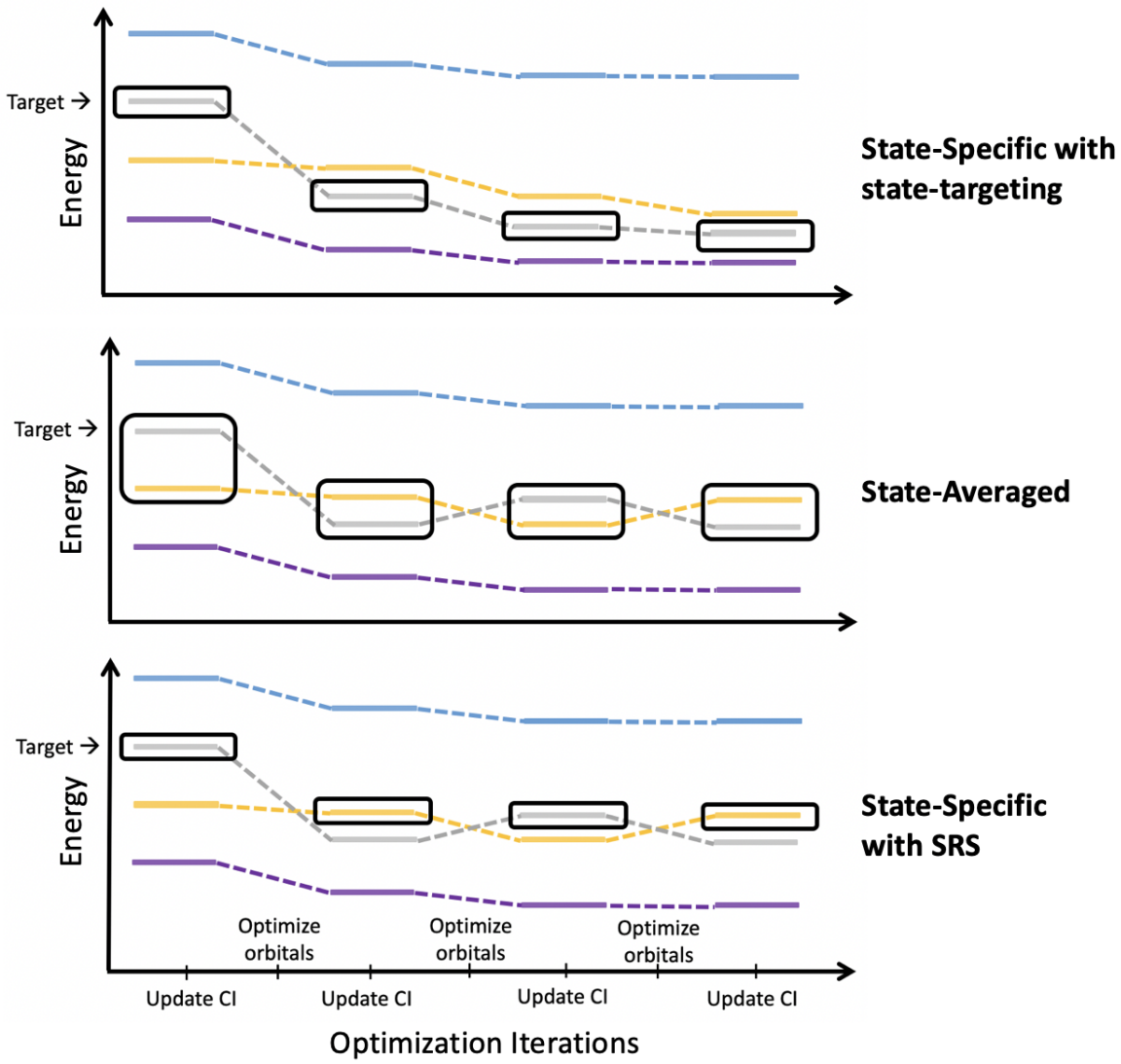


Figure 1.6. Optimization trajectories for various state-targeting approaches. The energy ordering of four states is plotted through iterative updates to the CI coefficients and relaxation of the orbitals where the third state (gray) is the target of the optimization.

1.8.3.1 Simple Root Selection

Root flipping is particularly problematic when using what might be called "simple root selection" (SRS) for the overall two-step optimization. In SRS, once a state is selected, the molecular orbitals of whole system are optimized for that particular state. With the orbitals updated, the active space CI Hamiltonian is re-diagonalized. This state-specific process—depicted in the bottom panel of Figure 1.6—is repeated until the change in energy and the orbital and CI energy gradients are below a given threshold. In a ground state calculation, SRS is very straightforward as the targeted state will always be the one lowest in energy. For excited states, the root flipping problem is often seen. For example, as shown in Figure 1.6, when the orbital optimization targeting the second excited state lowered the energy of this state below what was previously the first excited, the energy hierarchy is now different and SRS is naively targeting the wrong state. Root flipping often leads to convergence issues where the two states perpetually oscillate back and forth or it results in total optimization failure.[16]

1.8.3.2 State Averaging

The most common remedy to the root flipping problem is state-averaging (SA), shown in the middle panel of Figure 1.6. Instead of targeting a single state, this approach minimizes the average energy of multiple states.[111] This will account for any root flipping that occurs during the optimization, provided it is between states that are included in the average. State averaging comes with many other advantages as well. Mixing in a small amount of the ground state can greatly stabilize the optimization convergence when targeting an excited state. State averaging is useful in describing non-adiabatic interactions like spin-orbit coupling[112] or vibronic coupling[113] between excited states and provides a balanced description of electronic structure in systems with strong non-adiabatic couplings.[114, 115, 116, 117] Additionally, it is often convenient to have one set of orbitals that describe multiple states in part because those states will be exactly orthogonal, which is not the case in state-specific approaches.

While state averaging is widely used in CASSCF optimizations, there are several limitations to this approach. None of the resulting wave functions are true energy stationary points. Consequently, this adds significant complexity for the nuclear gradient evaluations necessary when performing a geometry optimization.[118, 119] Moreover, the orbitals are optimized to describe an average of the states making them less accurate for any one state in the average. This effect is magnified when the states included in the average have significantly different character, where treating only their average can lead to a poor description of each individual state.

An example that highlights the inadequacies of state averaging is the excitation of 4-amino-benzonitrile (ABN) to the charge transfer surface. Upon excitation the planar geometry relaxes and the question is—does the molecule stay planar or does the pi bond twist in response? Experimental results show a twisted structure after excitation.[120] A state-

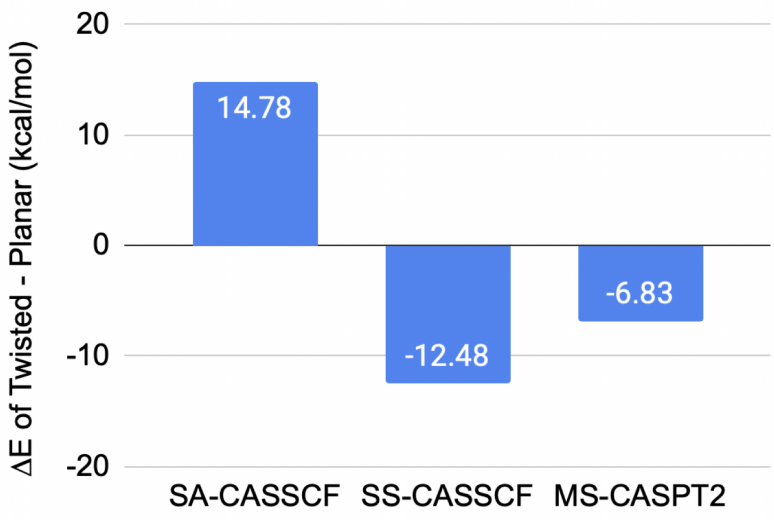


Figure 1.7. Bar chart of the SS, SA, CASPT2 energy twisted versus planar energy difference on the intramolecular charge-transfer surface of 4-aminobenzonitrile using a (12e,11o) active space with cc-pVDZ. Data values are taken from Tran and Neuscamman. [121]

averaging study however, found the planar structure to be almost 15 kcal/mol more stable on the charge-transfer surface than the twisted structure (Figure 1.7).[121] On the other hand, a state-specific CASSCF approach was found to agree with experiment. This example clearly demonstrates the qualitative inaccuracy of the state-averaged result due to its inability to fully relax orbitals for the CT state. The CT's substantial change to the molecule's dipole moment should be followed by a repolarization of the occupied orbitals' shapes, but the inclusion of the ground state in the state average works to resist this change as such repolarizations raise its energy. This is an example of the type of system and excitation where state-specific optimization succeeds where state-averaged does not.

1.8.3.3 State Targeting

Early evidence shows that the state-specific regime is preferable and even necessary for accurate predictions in systems with charge transfer excitations. If SRS cannot be relied on to bring a state-specific optimization to a healthy convergence, other approaches need to be considered to track the targeted state through the optimization and any potential root flipping. The top panel of Figure 1.6 shows the ideal scenario for a state-specific approach, where the orbitals are relaxed for the desired state at every point of the optimization regardless of changes to the energetic hierarchy.

Just as they have long been used in Δ-SCF to select orbitals similar to those in previous iterations,[65] MOM-like criteria can be used to track a CASSCF state across orbital updates by seeking to maximize the overlap between the successive wave functions. However, this

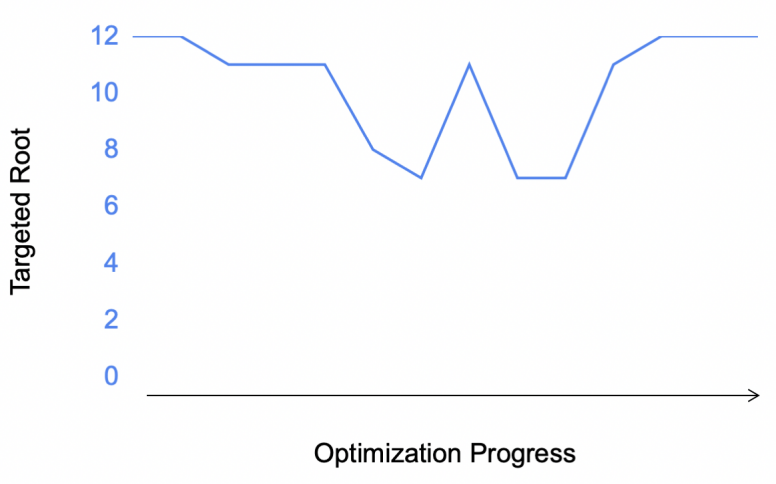


Figure 1.8. Targeted root across a small subset of the optimization steps during a calculation of a 1.8 Å MgO charge-transfer state 7^1A_1 using $W\Gamma$ with (8e,8o) and cc-pVDZ.

approach is not fully immune to root flipping issues. While there have been several variations of this published in the last few years, all approaches seem to struggle when a single step makes large updates to the wave function.[66, 67] The $W\Gamma$ method instead tracks the state by defining a quality measure based on approximate variational principles and overlap of the 1-body density matrices.[16] This approach is more robust but is still not immune to root flipping. This becomes clear with a close inspection of one of the high-lying charge transfer $^1A_1$ states of MgO (treated with 8 electrons in 8 orbitals with a cc-pVDZ basis set). Looking at the targeted state at each iteration for just a small subset of the optimization in Figure 1.8, it is clear there is significant root flipping occurring. This calculation did converge and, absent other information, one might hope that it converged to the correct state. However, the final wave function overlap with the initial CASCI root suggests otherwise. The final stationary point has only a 0.3 overlap where we would expect a successful calculation to be closer to an overlap of 1 as the initial CASCI, although imperfect, is a qualitatively correct approximation of the desired state. While MOM and $W\Gamma$ have made significant improvements state-specific optimizations for challenging scenarios like root-flipping, there are considerable opportunities for the one-step quasi-Newton methods discussed in this thesis to further improve state-specific optimization.

1.8.4 Post-CASSCF

CASSCF is routinely utilized as a foundational method for capturing the strong correlation of multi-reference excited states. Dynamic correlation is typically recovered by applying a perturbative expansion to the CASSCF reference wave function, CASPT2,[57, 58] capturing details such as wavefunction cusps and van der Waals correlations that are essential for overall

accuracy.[28] The combination of CASSCF and CASPT2 methods is a powerful tool for the accurate and reliable treatment of strongly correlated systems across many contexts such as bond breaking, bond forming, transition metals, and in particular for excited states.[122, 123, 124] Accurately representing these properties are vital for applications like dynamics and multi-state coupling in surface hopping.[125, 126, 127, 128] However, the successful outcome of a CASPT2 calculation relies heavily on the quality of the reference wave function, inheriting any optimization issues faced by CASSCF. Despite its successes, there remain important types of excited states, in particular states with significant CT character, that the traditional CASSCF approach struggles with. Early work in excited-state-specific CASSCF has shown promise in addressing this issue,[16, 121] but has not fully overcome the challenge of root flipping.

1.9 Outline

This thesis is the culmination of work on a fully excited-state-specific complete active space self-consistent field theory. We develop a GVP-based one-step optimization algorithm that overcomes the limitations of existing methods, demonstrating that this novel approach achieves superior accuracy, precision, and numerical stability.

Chapter 2 details the development of the GVP-CASSCF method and its application on a set of small molecules that display an array of optimization challenges. This work has been published in the article "Applying generalized variational principles to excited-state-specific complete active space self-consistent field theory" *J. Chem. Theory Comput.* 18, 6608 (2022).

Chapter 3 explores the capabilities of GVP-CASSCF in producing potential energy surfaces for excited states. We look closely at a set of excited states of thioacrolein with doubly excited and charge transfer character and analyze how the state-specific approach compares to state-averaging. We find that GVP-CASSCF can be configured to find either diabatic or adiabatic energy stationary points and that its adiabatic PES can be qualitatively different that produced by state averaging. As in ABN, these differences appear to be driven by differing abilities to capture post-CT orbital relaxations.

Finally, Chapter 4 summarizes the work presented in earlier chapters, discussing the advancements and impact of this work on excited state electronic structure theory of strongly correlated systems.

Chapter 2

Applying generalized variational principles to excited-state-specific CASSCF theory

2.1 Abstract

We employ a generalized variational principle to improve the stability, reliability, and precision of fully excited-state-specific complete active space self-consistent field theory. Compared to previous approaches that similarly seek to tailor this ansatz's orbitals and configuration interaction expansion for an individual excited state, we find the present approach to be more resistant to root flipping and better at achieving tight convergence to an energy stationary point. Unlike state-averaging, this approach allows orbital shapes to be optimal for individual excited states, which is especially important for charge transfer states and some doubly excited states. We demonstrate the convergence and state-targeting abilities of this method in LiH, ozone, and MgO, showing in the latter that it is capable of finding three excited state energy stationary points that no previous method has been able to locate.

2.2 Introduction

Whether one looks at carotenoids, [3, 129, 130] photochemical isomerization, [131, 132, 133] or transition metal oxide diatomics, [4, 134, 135] molecular excited states often display wave function characteristics that go beyond the simplifying assumptions of mean field theory. From the right perspective, this fact is not that surprising, as it is the widening of the HOMO-LUMO gap that helps determine ground state equilibrium geometries and ensure the validity of mean field theory. Upon excitation, a molecule may be far from the excited state's equilibrium geometry, and in any case there is no longer the HOMO-LUMO gap to prevent near-degeneracies between different fillings of the molecular orbital diagram that may be important for the state under study. The result is that methods like time-dependent

density functional theory and equation-of-motion coupled cluster theory that perturb around the mean field limit, while extremely useful in many excited state contexts, are qualitatively inappropriate in many others. Instead, methods that explicitly engage with the strongly multi-configurational nature of these states are called for. Ideally, these methods would be equally capable for excited states as they are for ground states, but, as in so many areas of electronic structure theory, the current reality is that they are not.

For decades, multi-configurational photochemical investigations have been supported by complete active space self consistent field (CASSCF) theory,[48, 49, 50, 51] but the approximations introduced in its most common incarnations can cause challenges when treating high-lying states or states with widely varying characters. In particular, the state averaging (SA) approach – in which one finds the orbitals that minimize the average energy of multiple configuration interaction (CI) roots – makes the assumption that all states of interest can be constructed to a similar degree of accuracy with one shared set of orbitals.[111] This approximation offers important advantages and has long been a standard and successful approach to excited states in CASSCF,[136, 114, 115, 137, 116, 117] but it can also create a number of difficulties. Most obviously, it is less appropriate in cases where different states require significantly different orbital relaxations, as occurs in molecules bearing both local and charge transfer (CT) excitations. Indeed, SA-CASSCF relative energies during nuclear motion on an charge transfer excitations' surface can be in error by 10 kcal/mol or more.[121] Further, the state averaging method links all of the states together so that if one state is not well served by the chosen active space and displays a non-analytic point on its energy surface, all states, even those well-served by the active space, will show cusps or discontinuities on their energy surfaces. Finally, because it is only the average energy that is made stationary with respect to the wave function variables, evaluating nuclear energy gradients for geometry optimization or dynamics requires solving difficult response equations which are indeed approximated in some implementations.[118, 119] In ground state CASSCF, by contrast, the state's energy is stationary already and nuclear gradient evaluations are much more straightforward. So, although state averaging has been and will continue to be a powerful asset to quantum chemical investigation, there are many reasons why and many settings in which a fully excited-state-specific CASSCF would be valuable.

Looking at the wider world of excited state theory, there has been remarkable progress in formulating fully state-specific methods in recent years, which augurs well for progress in this direction in CASSCF theory. Examples of this progress include work in variational Monte Carlo, [138, 19, 139] variance-based self-consistent field (SCF) theory, [140, 141] more robust level shifting approaches in SCF methods, [142] core spectroscopy, [143, 144, 145, 146] perturbation theory, [106] and coupled-cluster theory. [147, 148] Especially relevant to the current study is the "WΓ" approach to state-specific CASSCF (SS-CASSCF), [16] in which an approximate variational principle and density matrix information are used to carefully follow a particular CI root during a two-step optimization that goes back and forth between orbital relaxation steps and CI diagonalization steps. The WΓ approach proved capable of overcoming root flipping in a wider variety of situations than readily-available alternatives, improving CASPT2 energies when compared to state-averaging, and in making qualitative

improvements to some potential energy surfaces.[16, 121] However, it was unable to locate at least one of the low-lying states of MgO and, as a method that lacks coupling between orbital and CI variables, it struggles to tightly converge stationary points. The method presented here proves more reliable when faced with root flipping and far superior at tight convergence thanks to its objective function and its coupling of orbital and CI parameters during optimization.

To understand how these advantages come about, let us turn to discussing recent progress in the use of quasi-Newton methods to minimize energy-gradient-based objective functions, which has proven effective in the context of both the excited state mean field (ESMF) ansatz [105, 21, 60] and Kohn-Sham ΔSCF. [73] Essentially, the idea is to search for energy saddle points – which in full CI (FCI) would be the exact excited states – by minimizing the norm of the energy gradient with respect to the variational parameters. By relying on either an initial guess sufficiently close to the desired stationary point [73] or a generalized variational principle (GVP) that can use sought-after properties to steer an optimization towards that stationary point, [21] these approaches have proven capable of achieving full excited-state-specificity while avoiding root flipping or variational collapse to lower states. While the work in this direction so far has mostly been focused on weakly correlated excited states, there is no formal barrier to applying the GVP approach to the CASSCF ansatz, which is our focus here.

To perform excited-state-specific optimization of the CASSCF ansatz, we will minimize a GVP containing the square gradient norm by purely quasi-Newton descent, eschewing CI diagonalization (except in generating a guess) and more traditional augmented Hessian approaches to orbital rotations.[149, 150] Of course, it may be that a combination of all of these methods ultimately proves more efficient, as has recently been found for the ground state,[17, 84, 85] but in this first combination of CASSCF with a GVP, we stick to pure quasi-Newton minimization for simplicity, and so our core computational task is to evaluate gradients of an objective function that contains the square norm of the energy gradient. Recent work has provided multiple ways forward here. On the one hand, automatic differentiation arguments guarantee that in most scenarios, the requisite derivatives can be derived automatically and will have a cost that is a modest and constant multiple of the energy evaluation cost.[105] In many cases, this guarantee can motivate the derivation of analytic forms for these derivatives,[109] which are often even more efficient in practice, although not necessarily simple or easy to implement. As an alternative, Hait and Head-Gordon have presented a clever finite-difference approach to these derivatives.[73] Although finite difference will incur some error relative to analytic or automatic differentiation, their study of orbital optimization shows that this error is small enough that it does not prevent successful convergence to excited state stationary points. The key benefit of this approach is that it requires only that the energy gradient itself be available, and so is more convenient to implement. Although it is possible that a fully analytic formulation of the energy gradient norm derivatives would improve the rate of quasi-Newton convergence by avoiding finite difference errors, we for simplicity adopt the finite difference approach here and find that optimization remains effective even when orbital and CI parameters are optimized together. In future, it may be interesting to ex-

plore whether more accurate analytic expressions improve numerical efficiency and whether mixtures with CI and augmented Hessian orbital optimizers are worthwhile, but already the present approach to combining CASSCF with an excited state GVP allows us to succeed in situations where previous CASSCF approaches fail.

2.3 Theory

2.3.1 CASSCF Ansatz

The standard CASSCF ansatz[48, 49, 50, 51] has been the foundation for a wide range of CASSCF derived methods,[75, 76, 52, 77, 53, 78, 79, 85] and is the formulation used in the approach introduced here. CASSCF methods classify subsets of the molecular orbitals as closed orbitals each occupied by two electrons, active orbitals with varying occupation, and virtual orbitals that are completely unoccupied. The CASSCF wave function is therefore composed of all possible electronic configurations within the active orbitals, defining the active space. The wave function must also account for orbital relaxation effects as while rotations within the active space are described entirely by changes to the configuration (CI) coefficients, the virtual and closed orbitals remain excluded. While enlarging the active space captures more orbital relaxation effects via the CI expansion, this quickly becomes computationally infeasible for large systems. In addition, the results of a CASSCF calculation are often used as the input for higher-order methods that recover dynamic correlation, which can further limit the size of the chosen active space. Instead, to relax the orbital descriptions we incorporate an orbital rotation operator in the wave function, such that

$$|\Psi_{\mathrm{CAS}}\rangle = e^{\hat{X}} \sum_{I} c_I \, |\phi_I\rangle \tag{2.1}$$

where $|\phi_I\rangle$ represents a Slater determinant and c_I is the corresponding CI coefficient. The total number of Slater determinants, and thus CI variational parameters forming $\vec{c}$, is determined by the size of the active space.

For a finite basis of spatial orbitals, the operator $\hat{X}$ in Eq. (2.1) is given by

$$\hat{X} = \sum_{p<q}^{N_{\mathrm{basis}}} X_{pq} \left(\hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p \right). \tag{2.2}$$

It is defined to be real and spin restricted, thereby ensuring the orbital rotation operator $\hat{U} = e^{\hat{X}}$ is unitary and also spin restricted. [29, 21] Note that only the upper triangle of the matrix $\boldsymbol{X}$ appears in Eq. (2.2), although it is often useful to consider the full matrix, which is anti-Hermitian and thus defined by the upper triangle. Additionally, rotations between orbitals within the active space do not affect the energy as they are redundant with the flexibility present in the CI expansion. Similarly, rotations within the closed and virtual orbital spaces have no affect on the energy. Were these redundant parameters retained, the variable space

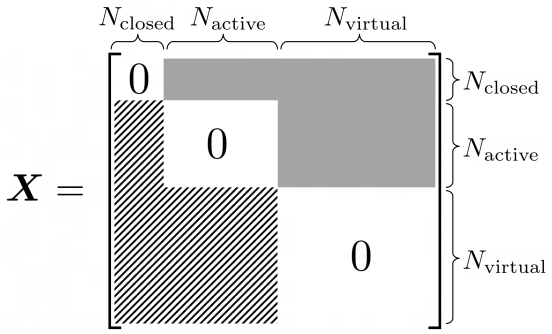


Figure 2.1. Orbital rotation coefficient matrix $\boldsymbol{X}$ where the solid shaded area represents nonzero variational parameters, and the striped region is the negative transpose.

would contain an infinite seam of energetic degeneracy, and so to avoid complications during numerical optimization, all redundant parameters are excluded. This choice leads to Figure 2.1, which shows the blocks of $\boldsymbol{X}$ that are included in the orbital variational parameter set $\vec{x}$. All together, our CASSCF wave function's variational parameters are the concatenated set $\vec{v} = \{\vec{c}, \vec{x}\}$.

2.3.2 Objective Function

2.3.2.1 Generalized Variational Principle

In FCI, when the energy is expressed as a function of the CI coefficients, the exact excited states are the energy saddle points of this function. Even in more approximate theories, the approximate ansatz's saddle points are often good approximations to the excited states,[65, 66, 67, 105] and thus the focus of the present investigation is to find excited state energy stationary points for the CASSCF ansatz. As these points are not energy minima, gradient-based descent methods are likely to collapse to lower states, and even non-gradient-based methods like self-consistent field algorithms can display similar difficulties.[66, 67] To retain the convenience of minimization algorithms while avoiding this issue of variational collapse, we choose objective functions that have the square norm of the energy gradient as their centerpiece.

$$|\nabla_{\vec{v}} E|^2 = \sum_i \left|\frac{\partial E}{\partial c_i}\right|^2 + \sum_j \left|\frac{\partial E}{\partial x_j}\right|^2 \tag{2.3}$$

In CASSCF, this gradient norm contains contributions from both the CI coefficient gradients and the orbital rotation gradients. It is positive semi-definite by construction, and, for an isolated energy saddle point, is expected to be surrounded by a basin of convergence that, if we can somehow get ourselves inside it, should allow a straightforward minimization of $|\nabla_{\vec{v}} E|^2$ to bring us to the desired excited state energy stationary point. It is important to note

that when $\nabla_{\vec{v}}|\nabla_{\vec{v}}E|^2 = 0$ it is possible that $|\nabla_{\vec{v}}E|^2 \neq 0$, meaning that the square gradient norm has stationary points that are not energy stationary points. In the results discussed below, such cases were overcome through a combination of improved initial orbital guesses and by incorporating additional properties within the generalized variational principle [21] (GVP) to which we now turn our attention.

With the norm of the energy gradient being zero for all energy stationary points, we require some mechanism by which the desired excited state's stationary point can be targeted. In some cases, a good enough guess is available to place one within the appropriate basin of convergence, but in general such a guess may not be available. To address this problem, we use a GVP approach to expand our objective function beyond the square gradient norm so that other properties of the excited state can help steer the optimization into the desired convergence basin.

$$L_\mu = \mu \left|\vec{d}\right|^2 + (1-\mu)\left|\nabla_{\vec{v}}E\right|^2 \tag{2.4}$$

In this objective function, $\vec{d}$ contains functions of the wave function that should have values close to zero for the desired excited state, such as the difference $\langle\hat{H}\rangle - \omega$ between the current wave function energy and a guess for the excited state's energy. Thus, when μ is greater than zero and we minimize L_μ, the term containing $\vec{d}$ should help drive the optimization towards the energy stationary point belonging to the desired excited state. If the functions within $\vec{d}$ uniquely specify the state (by which we mean the norm of $\vec{d}$ is smaller for that excited state than for any other energy stationary point), then an optimization in which μ is gradually lowered to zero will arrive at the desired stationary point. [21]

The energy difference term $\langle\hat{H}\rangle - \omega$ that we typically include within $\vec{d}$ can be motivated as a useful approximation [105, 16] to the rigorous excited state variational principle

$$W = \frac{\langle\Psi|\,(\omega-\hat{H})^2\,|\Psi\rangle}{\langle\Psi|\Psi\rangle} \approx \left(\langle\hat{H}\rangle - \omega\right)^2, \tag{2.5}$$

which if evaluated exactly has its global minimum at the Hamiltonian eigenstate whose energy is closest to ω. [71, 72] Of course, many other properties and functions of the wave function can also be useful in specifying the desired state through the vector $\vec{d}$. For example, if we knew that it should ideally be orthogonal to another nearby state $|\Phi\rangle$ and should have a dipole moment $\vec{\mu}$ (not to be confused with the weighted average parameter μ above) of about $\vec{\mu}_0$, we might use $\vec{d} = \{\langle\hat{H}\rangle - \omega,\, \langle\Psi|\Phi\rangle\,,\, |\vec{\mu} - \vec{\mu}_0|\,\}$ to guide our optimization into the desired basin of convergence, at which point μ can be reduced to zero so that, in the final stage of optimization, minimization of the energy gradient square norm brings us to the desired stationary point. It is important to recognize that the functions employed within $\vec{d}$ need not be exact, as their only purpose is to get us into the right basin of convergence, after which they have no further effect. A good example of where this flexibility can be exploited is seen in our results on ozone, where we use a simple approximation for the overlap with another state to help one of our optimizations converge correctly. Evaluating that overlap

exactly would be an exercise in non-orthogonal CI (NOCI),[151, 152, 153] but in this case a simple dot product between CI vectors (which neglects differences in the molecular orbitals) is free by comparison and a good enough nudge to guide the optimization to the desired stationary point in the face of a tricky near-degeneracy.

2.3.2.2 Objective Function Gradient

To minimize our objective function via gradient descent, we will need an expression for its gradient. When $\vec{d} = \{\langle \hat{H} \rangle - \omega\}$, this gradient is

$$\nabla_{\vec{v}} L_\mu \;=\; 2\mu(E-\omega)\nabla_{\vec{v}} E \;+\; (1-\mu)\nabla_{\vec{v}} |\nabla_{\vec{v}} E|^2. \tag{2.6}$$

In CASSCF, the energy gradient with respect to the full variational parameter set $\nabla_{\vec{v}} E$ can be split into the energy gradient with respect to the CI parameters $\nabla_{\vec{c}} E$ and the energy gradient with respect to the orbital rotation parameters $\nabla_{\vec{x}} E$. In this work, we use the analytic expression for the CI gradient

$$\nabla_{\vec{c}} E = \frac{\partial E}{\partial \vec{c}} = \frac{2(H-E)\vec{c}}{\vec{c}^T \cdot \vec{c}} \tag{2.7}$$

where H is the Hamiltonian matrix in the CI basis. For the orbital energy gradient, we use the analytic expressions given in Appendix A that are comprised of contractions between the MO integrals and the one and two-electron spin-summed reduced density matrices. [109, 140, 73, 49, 51, 24] These expressions assume we are working within the current MO basis (i.e. when $X = 0$), the implications of which are discussed in Section 2.3.2.3.

By far the most computationally challenging term in Eq. (2.6) is the derivative of the squared norm of the energy gradient with respect to the variational parameters,

$$\frac{\partial}{\partial v_j} |\nabla_{\vec{v}} E|^2 \;=\; \frac{\partial}{\partial v_j} \sum_i \left| \frac{\partial E}{\partial v_i} \right|^2 \;=\; 2 \sum_i \mathcal{H}_{ij} \frac{\partial E}{\partial v_i}. \tag{2.8}$$

The Hessian matrix of energy second derivatives $\mathcal{H}_{ij} \equiv \frac{\partial^2 E}{\partial v_i \partial v_j}$ is expensive to evaluate, and we certainly do not wish to construct it explicitly. While it is possible to use automatic differentiation to evaluate this term, [105] for ease of implementation we instead turn to a central finite difference method that Hait and Head-Gordon have shown to be effective for excited state orbital optimization. [73] Using a directional finite difference of the energy gradient with a chosen perturbation of $\delta\vec{v} = \lambda \nabla_{\vec{v}} E\big|_{\vec{v}=\vec{v_0}}$ yields the approximate expression

$$\nabla_{\vec{v}} |\nabla_{\vec{v}} E|^2 = \frac{1}{\lambda} \left(\nabla_{\vec{v}} E\big|_{\vec{v}=\vec{v_0}+\delta\vec{v}} - \nabla_{\vec{v}} E\big|_{\vec{v}=\vec{v_0}-\delta\vec{v}} \right) + O\left(\lambda^2 \left(\nabla_{\vec{v}} E\big|_{\vec{v}=\vec{v_0}} \right)^3 \right). \tag{2.9}$$

This approach avoids the computationally demanding Hessian-gradient contraction in Eq. (2.8), replacing it with multiple evaluations of the energy gradient. Automatic differentiation – as its cost is typically 2-3 times the cost of the function – should be able to deliver a

fully analytic version of this approach with zero finite difference error at a similar price, as has been achieved for ESMF. Further, a hand-implemented analytic version could be even faster. Thus, it may be worth investigating in future whether the removal of the small finite difference error leads to a significant improvement in optimization efficiency. For the present study, however, we employ Eq. (2.9) as is for both the orbital and CI variables together and find that it is sufficient for achieving tight energy stationary point convergence. It is important to stress that, regardless of which of these approaches is taken for evaluating the objective function gradient, the computational cost of doing so is at worst equal to a handful of CASSCF energy gradient evaluations, and so the scaling of the approach with system size is the same as in standard CASSCF.

A close inspection of Eq. (2.8) shows that, even if one applies naive steepest descent for minimizing the objective function, some coupling between the orbital and CI variables is present due to the energy Hessian. In practice, a quasi-Newton approach that builds up an approximation to the objective function Hessian will account for even more coupling between these variable sets. Although it is too early to tell how well this approach to coupling works as compared to second-order ground state approaches, [17, 85] a quasi-Newton minimization of our objective function certainly incorporates more coupling than a simple two-step optimization [16] in which one goes back and forth between optimizing the CI variables with the orbitals held fixed and optimizing the orbitals with the CI variables held fixed. In each step of quasi-Newton minimization, the effects of orbital changes on the CI energy gradient and CI changes on the orbital energy gradient are taken approximately into account. The result is a dramatic improvement in the method's ability to tightly converge the energy gradient as compared to the two-step WΓ approach that we compare to in our results below.

2.3.2.3 Approximate Objective Function Hessian

In this work, we use the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [154, 155, 156, 157] to minimize the objective function. Roughly speaking, L-BFGS takes a Newton-like step using an approximate Hessian. In particular, this approximate Hessian is arrived at by using finite-differences between previous iterations' objective function gradients to improve upon some initial guess for the objective function Hessian. This initial guess can be set to the identity matrix for simplicity, but the speed of convergence can be accelerated dramatically if a better guess for the Hessian is supplied[82], as has been demonstrated for objective functions like ours in both the ΔSCF [65] and ESMF [158] contexts. Indeed, our approach here is another example of using a quasi-Newton method to further improve a CASSCF approximate Hessian scheme. An early example of using quasi-Newton methods for this purpose occurred in the context of improving super-CI methodology for restricted active space wave functions, [159] and very recent work has shown that orbital-CI coupling for ground state optimizations can be usefully accelerated via quasi-Newton as well. [17, 84] In the present study, we see that even if L-BFGS starts from the identity matrix as the initial Hessian guess, it is better at achieving tight convergence than an uncoupled

two-step optimization like the $W\Gamma$ method. However, the smarter approach [159, 17, 84] of using a quasi-Newton method like L-BFGS to improve on a more accurate (although still approximate) initial Hessian is more effective still, and so we will seed L-BFGS with diagonal approximations to our objective function's Hessian.

Starting with the Hessian of the $\mu = 0$ objective function, (i.e. the second derivatives of the energy gradient square norm)

$$\frac{\partial^2}{\partial v_j \partial v_k} \sum_i \left| \frac{\partial E}{\partial v_i} \right|^2 = 2 \sum_i \mathcal{H}_{ij} \mathcal{H}_{ik} + 2 \sum_i \left(\frac{\partial^3 E}{\partial v_i \partial v_j \partial v_k} \right) \frac{\partial E}{\partial v_i}, \tag{2.10}$$

we can anticipate that, due to its contraction with the energy gradient, the role of the third derivative tensor will become negligible as the optimization approaches an energy stationary point. Indeed, it has been observed empirically in both ΔSCF [65] and ESMF [158] that dropping this term entirely does not much matter, and so we neglect it here as well. In the case where $\vec{d} = \{\langle \hat{H} \rangle - \omega\}$ and we now allow μ to be zero or nonzero, this leaves us with the following approximate expression for the objective function Hessian.

$$\frac{\partial^2 L_\mu}{\partial v_j \partial v_k} \approx 2\mu \left[(E - \omega) \mathcal{H}_{jk} + \frac{\partial E}{\partial v_j} \frac{\partial E}{\partial v_k} \right] + 2(1 - \mu) \sum_i \mathcal{H}_{ij} \mathcal{H}_{ik} \tag{2.11}$$

When not using the identity, we will use the diagonal of Eq. (2.11) as the approximate objective function Hessian that we supply to L-BFGS. However, evaluating the full energy Hessian $\mathcal{H}$ is impractically expensive. To make this approach affordable, we extend the diagonal approximation to $\mathcal{H}$ as well, leaving us with the following expression.

$$\frac{\partial^2 L_\mu}{\partial v_i^2} \approx 2\mu \left[(E - \omega) \mathcal{H}_{ii} + \left| \frac{\partial E}{\partial v_i} \right|^2 \right] + 2(1 - \mu) \mathcal{H}_{ii}^2 \tag{2.12}$$

We approximate the energy Hessian $\mathcal{H}$ in Eq. (2.12) using a diagonal form, although we make different choices for how to deal with the CI block (denoted ${}^{cc}\mathcal{H}$) and the orbital block (denoted ${}^{xx}\mathcal{H}$). In the CI block, we make no approximation beyond omitting the off-diagonal terms, leaving us with the same diagonal that is used in the Davidson algorithm. [81]

$${}^{cc}\mathcal{H}_{ii} = \frac{2\,(H_{ii} - E)}{\vec{c} \cdot \vec{c}} \tag{2.13}$$

For the diagonal of the orbital block, we define $E_{pq}^- = \left(\hat{a}_p^\dagger \hat{a}_q - \hat{a}_q^\dagger \hat{a}_p \right)$ and arrive at the following expression. [29]

$${}^{xx}\mathcal{H}_{pq,pq} = \frac{\partial^2 E}{\partial x_{pq} \partial x_{pq}} = \langle \Psi | \left[E_{pq}^-, \left[E_{pq}^-, \hat{H} \right] \right] | \Psi \rangle \tag{2.14}$$

Following the derivation by Siegbahn *et al.* of the full orbital-orbital energy Hessian using Fock-like matrices,[83, 78] explicit expressions for the exact diagonal of ${}^{xx}\mathcal{H}$ in terms

of two-electron integrals and density matrices are provided in Appendix A for the reader's convenience and have been extensively checked with finite difference. Previous approaches in second-order MCSCF methods make further approximations to the diagonal of ${}^{xx}\mathcal{H}$, demonstrating this to be sufficient to achieve improved convergence.[160, 76] In addition to implementing the exact diagonal expressions and unlike the CI block, we go beyond just dropping the off-diagonal terms by approximating the Hamiltonian inside the commutators with the one-electron Fock operator built from our CASSCF wave function's one-body density matrix. These choices for our approximate energy Hessian diagonal, which are similar to those made in other contexts, [65, 158] combine with Eq. (2.12) to provide L-BFGS with a much better guess than the identity for the objective function Hessian. The Fock-based diagonal improved guess comes at an additional computational cost that is significantly less than the energy gradient evaluation we are already doing, as it involves no two-electron AO-to-MO integral transforms and has a much simpler interaction with the CI vector. While the exact diagonal version necessitates additional AO-to-MO integral transforms not already performed, for the small molecules considered in this study we find the increased cost to be off-set by the convergence speed-up it offers.

In practice, the working equations for the gradients and Hessian elements we need are simpler when the orbital rotation matrix X is equal to zero, as it is at the start of the optimization. However, if one uses the straightforward parameterization of the ith iteration's molecular orbitals as a single rotation from the initial guess,

$$C_i = C_0 e^X \tag{2.15}$$

then at all iterations aside from the first, one must deal with a non-zero X matrix. If, instead, one resets the definition of the molecular orbitals so that X becomes the rotation from the previous iteration's orbitals

$$C_i = \tilde{C} e^X = C_0 e^{X_1} e^{X_2} ... e^{X_{i-1}} e^X \tag{2.16}$$

then the working equations at each iteration enjoy the simplicity offered by having $X = 0$. However, when we reset the definition of X in this way, we cause the gradient history we have accrued to no longer be quite correct, as those gradients were evaluated with a slightly different definition of the variables. In previous work on single-determinant wave functions, [74] it has been shown that the gradient history can be exactly corrected to account for this change of variables. For simplicity, we have not done so here, and this has not prevented our approach from achieving tight convergence for excited states. However, making these types of gradient history corrections will presumably accelerate our rate of convergence, and so we look forward to investigating these corrections in future efforts to improve numerical efficiency, which could also benefit from the use of more sophisticated initial Hessians with non-zero orbital-CI blocks.

2.3.3 Optimization Procedure

The overall quasi-Newton optimization procedure for our GVP approach to excited state CASSCF is as follows.

1. An initial orbital basis and active space are chosen and an initial guess for the CI coefficients is selected, typically taken from a CASCI calculation or an initial SA-CASSCF calculation. The orbital rotation coefficients are initialized as zero and a value for ω is estimated from the energy of the initial inputs, results from other methods, or experimental data.

2. The set of variational parameters $\vec{v} = \{\vec{c}, \vec{x}\}$ are optimized all together via a series of L-BFGS minimizations of L_μ for decreasing values of μ. We supply either the identity or an approximate objective function Hessian discussed in the previous section as the initial guess for the L-BFGS Hessian. The initial μ value and convergence threshold are set to 0.5 and $|\nabla_{\vec{v}} L| = 10^{-3}$, respectively. Within each micro-iteration of an L-BFGS minimization, the following tasks are completed.

   a) The gradient of the objective function $\nabla_v L_\mu$ with respect to the CI coefficients $\vec{c}$ is built from the analytical expression in Eq. (2.7) where the contraction of the active space Hamiltonian with the CI coefficient vector is performed utilizing PySCF's[161] existing direct CI functions.

   b) The gradient with respect to the orbital rotation coefficients $\vec{x}$ evaluated at $X = 0$ is built from Eq. (A8)-(A10). The scaling of this task is dominated by the AO-to-MO integral transformations.

   c) The value of the finite difference λ is set to the maximum of $\{10^{-6}, |\nabla_{\vec{v}} E|\}$ at each iteration, and the objective function (Eq. (2.4)) and its gradient (Eq. (2.6)) are built at the cost of three gradient evaluations of both $\nabla_{\vec{c}} E$ and $\nabla_{\vec{x}} E$.

   d) If the approximate objective function Hessian (Eq. (2.12)) is in use, then it is built using either the exact energy Hessian diagonal or its Fock-based approximation as discussed in the previous section.

   e) Take the L-BFGS step and, afterwards, update the definition of the MOs as discussed in the previous section so that $X = 0$ again.

3. After each L-BFGS minimization (macro-iteration), we reduce μ. If the maximum element of $|\nabla_{\vec{v}} E|$ is now less than the current convergence threshold, then we jump to the final optimization stage, setting $\mu = 0$ and the convergence threshold to its final value of 10^{-7} and repeating step 2. Otherwise, we decrease μ by 0.1 and tighten the convergence threshold by a factor of 10 (if it is not yet 10^{-7}) and repeat step 2.

2.4 Results and Discussion

In the following collection of molecular examples, we aim to answer the key question of how does the GVP approach compare to other SS-CASSCF methods? Is the GVP able to find the CASSCF energy stationary point that corresponds to the initial CASCI root in the face of root-flipping? How does the convergence of the GVP approach compare to other SS-CASSCF methods, with and without the approximate diagonal Hessian being provided to

L-BFGS? Finally, are there situations where the GVP can succeed when other SS-CASSCF methods fail?

These questions were investigated in LiH, asymmetrically stretched O_3, and MgO. The cc-pVDZ atomic orbital basis[162, 163] was used throughout. Both LiH and O_3 used the HF orbital basis for the initial guess, while MgO used the local density approximation (LDA) orbital basis. An initial CASCI calculation was performed for each of these molecules and the targeted root's CASCI CI vector was used as the initial guess for the CI coefficients. Values for ω were chosen using past results from other CASSCF calculations or estimated based on the initial CASCI energy orderings. The first macro-iteration of each GVP optimization performed in this study held the CI parameters fixed while converging the orbital gradient to $|\nabla_{\vec{x}} L| < 10^{-5}$, using the identity as the objective function Hessian guess. Beyond the first macro-iteration, all parameters were optimized together with the approximate diagonal Hessian guess — built from the exact diagonal energy Hessian — employed for all values of μ in all optimizations in LiH, O_3 and MgO.

In this study, we consider a stationary point converged in our GVP optimization when $|\nabla_{\vec{v}} |\nabla_{\vec{v}} E|^2| < 10^{-7}$, $|\nabla_{\vec{c}} E| < 10^{-6}$, and $|\nabla_{\vec{x}} E| < 10^{-6}$. For each of the molecules in this study, the results of the GVP approach are compared to those of the $W\Gamma$ and simple root selection (SRS) 2-step methods. In SRS, one selects the CI root to use in orbital optimization by always taking the nth root from the energy-ordered CI roots, whereas $W\Gamma$ uses an approximate variational principle and the one-body density matrix to select the desired root. [16] For both $W\Gamma$ and SRS, neither of which has orbital-CI coupling in our implementation, we set looser convergence thresholds because this lack of coupling prevents them from converging to the same level of precision. For the change in energy, the norm of the orbital gradient, and the norm of the change in the one-electron density matrix, the $W\Gamma$ thresholds were set to 10^{-7}, 10^{-4}, and 10^{-4} respectively. To check whether a loosely converged $W\Gamma$ or SRS calculation corresponds to the same stationary point as the GVP, we have therefore also used our GVP approach to finalize their convergence. This finalization was never observed to alter the character of the wave function, even in cases where a non-negligible energy change was observed during finalization. All molecular orbital analysis was performed with the programs Gabedit [164] and Molden. [165]

2.4.1 LiH

The ground state of LiH ($X^1\Sigma^+$) is ionic at its equilibrium bond length of 1.8 Å, but the first excited state ($A^1\Sigma^+$) is mostly neutral due to a HOMO-LUMO charge transfer excitation. However, as the bond is stretched, the ground state becomes increasingly neutral while the first excited state becomes more ionic. What makes this an especially interesting molecule to study in the present context is the avoided crossing that exists between the ground and first excited states at intermediate bond lengths. [37, 166] The mixing of state characters in this region leads to a well known root flipping problem [16, 21, 37, 166, 111, 167] that provides a good test for our GVP approach.

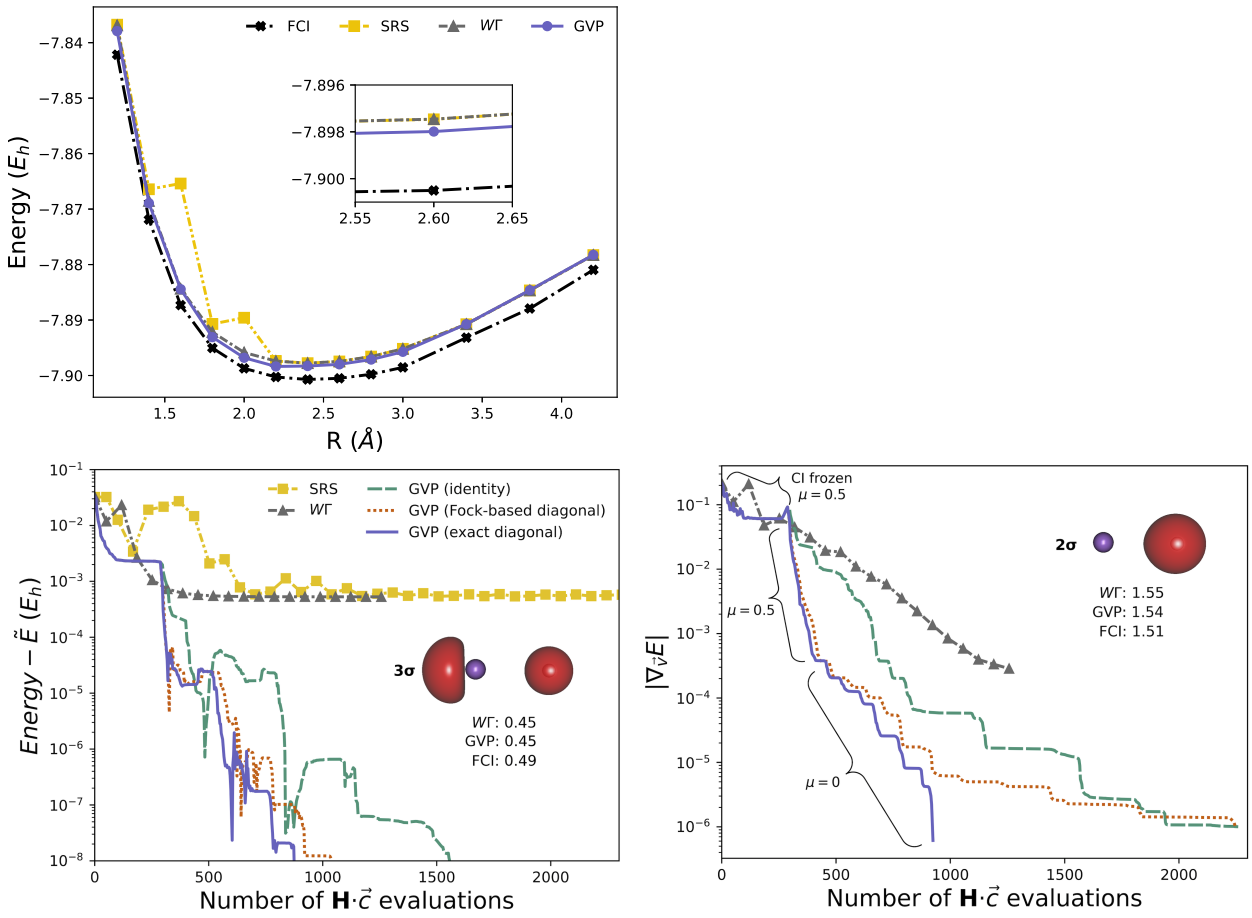


Figure 2.2. The top panel shows potential energy surfaces for the first excited state of LiH. The bottom left panel shows energy convergence at a bond length of 2.6 Å relative to the GVP's final tightly converged energy $\tilde{E}$. The bottom right panel shows, again at 2.6 Å, the convergence of the norm of the energy gradient. In both bottom panels, the optimization details are labeled for each macro-iteration of the GVP approach employing the μ update schedule as described in Section 2.3.3, with $\omega = -7.9$ E_h used at all macro-iterations. Convergence of the GVP is shown using the identity as the initial Hessian guess (dashed green line), compared to an approximate Hessian built from the exact diagonal (solid purple line) or Fock-based approximate diagonal (dotted orange line) energy Hessian. The insets to the bottom panels show the 2σ and 3σ natural orbitals and corresponding occupation numbers. At each geometry, SRS and $W\Gamma$ converged the orbital gradient to 10^{-4}, while the GVP converged to 10^{-7}.

Using an active space of 4 electrons in 4 orbitals (Li 1s2s2p$_z$, H 1s), Figure 2.2 demonstrates that SRS clearly suffers from the root flipping problem, causing it to struggle with convergence and taking a comparatively large number of iterations or failing altogether. Past work[16] has shown that the $W\Gamma$ method is able to overcome the root flipping problem by tracking the targeted root through the optimization, producing the smooth potential energy surface seen in the top panel of Figure 2.2. While the dissociation curves illustrate the agreement between $W\Gamma$ and the GVP approaches across all geometries, they also highlight the improvement the GVP achieves in overall convergence, in particular the magnitude of the orbital gradients, by several orders of magnitude from both the SRS and $W\Gamma$ results. For a geometry of 2.6 Å, Table 2.1 shows very similar wave function character between the energy stationary point the GVP finds and the more loosely converged $W\Gamma$ state. Both have strong overlap to the initial CASCI root and it is clear they are both describing the desired state, one is merely more tightly converged than the other. Indeed, looking at the convergence for this geometry in the bottom panels of Figure 2.2, the GVP achieves an energy half a mE$_h$ closer to the FCI result than the other state-specific methods in fewer Hamiltonian-CI-vector multiplies when using an approximate initial Hessian guess in L-BFGS. It is especially noteworthy that when using the identity as the initial Hessian guess we still take a comparable number of Hamiltonian-CI vector contractions, suggesting that helpful orbital-CI coupling is indeed present in the quasi-Newton approach even without the better Hessian starting guess.

Table 2.1. Wave function character in the CASCI orbital basis of the first excited state $A^1\Sigma^+$ of LiH at a bond length of 2.6 Å.

Primary Excitations	Active Space Electron Configuration	Wavefunction Weight (%) CASCI	$W\Gamma$	GVP
$2\sigma \rightarrow 3\sigma$	$1\sigma^2\ 2\sigma\ 3\sigma$	86.5	82.5	82.9
$2\sigma^2 \rightarrow 3\sigma^2$	$1\sigma^2\ 3\sigma^2$	5.3	5.7	5.7
$2\sigma^2 \rightarrow 3\sigma,\ 4\sigma$	$1\sigma^2\ 3\sigma\ 4\sigma$	4.2	5.6	5.6
Aufbau	$1\sigma^2\ 2\sigma^2$	3.2	5.3	4.9
Overlap with CASCI Root:		1	0.95	0.96

2.4.2 Asymmetrical O_3

We turn next to asymmetrically stretched ozone, which contains two excited states that are close to energetically degenerate and prove to be especially challenging for the GVP approach. Indeed, at this particular geometry ($R_{O_1O_2} = 1.3$ Å, $R_{O_2O_3} = 1.8$ Å, $\angle O_1O_2O_3 = 120°$), the 4^1A" and 5^1A" states can switch order with each other and even strongly re-mix their primary configurations depending on the size of the active space used and whether or not the orbitals are optimized state-specifically. We employ a 9-orbital, 12-electron active space and freeze

the electronic occupation and orbital shapes of the six lower energy orbitals (which are, roughly speaking, the O 1s and 2s orbitals). With this choice, we do in fact observe a root flip: SS-CASSCF optimizations starting from the 4th and 5th ^1A" CASCI roots find two different energy stationary points, but the stationary point found when starting from the 5th CASCI root (and which is most similar in character to the 5th CASCI root) has a lower energy than the other stationary point, as displayed in Table 2.2.

Table 2.2. Wave function data in the CASCI orbital basis for the 4th and 5th ^{1}A" states in O_3. Note that the 5th CASCI root ultimately optimizes to become the 4^1A" state, and so its data is presented under the 4^1A" heading in the left column, whereas the 4th CASCI root's data is presented on the right under the 5^1A" heading. The GVP data are for the stationary point found when starting from the CASCI root shown under the same heading.

	4^1A" Wavefunction Weight (%)		5^1A" Wavefunction Weight (%)	
Primary Excitations	CASCI	GVP	CASCI	GVP
9a', 10a' → 3a", 11a'	67.3	65.7	5.4	19.6
2a" → 11a'	1.8	6.0	41.0	40.1
9a', 2a" → $3a"^2$	1.4	0.0	10.0	4.0
Overlap with 4th ^{1}A" CASCI root	0	0.41	1	0.66
Overlap with 5th ^{1}A" CASCI root	1	0.87	0	0.68
Energy (E_h)	-224.258	-224.313	-224.265	-224.309

As seen in Figure 2.3, the initial CASCI states (when swapped in energy ordering) have very similar natural orbital occupation patterns as the SS-CASSCF energy stationary points, but a close inspection of the data in Table 2.2 suggests that the story is not entirely straightforward. Indeed, although the GVP optimization starting from the 5th CASCI root converges tightly and without incident to an energy stationary point, the final non-orthogonal-CI-style overlaps between this stationary point and the two CASCI roots (Table 2.2) show that a non-trivial remixing has occurred. The stationary point is still dominated by the CASCI root we started from (overlap 0.87), but contains a significant amount of the other root as well (overlap 0.41).

When attempting the GVP optimization starting from the 4th CASCI root, the story is even less straightforward, with our first attempt at minimizing the GVP failing to find a stationary point at all. While this difficulty eventually revealed itself to be an example of a bad initial wave function guess, this was not obvious until we had later found the 5^1A" stationary point and could verify that, indeed, the CASCI guess was pretty far from the mark. In practice, it will often be prudent to start from a better initial guess by using an equal or biased weighting in SA-CASSCF. Here, however, we intentionally keep this poor initial guess in order to investigate the efficacy of adding additional properties to the GVP to help guide the optimization into the correct basin of convergence.

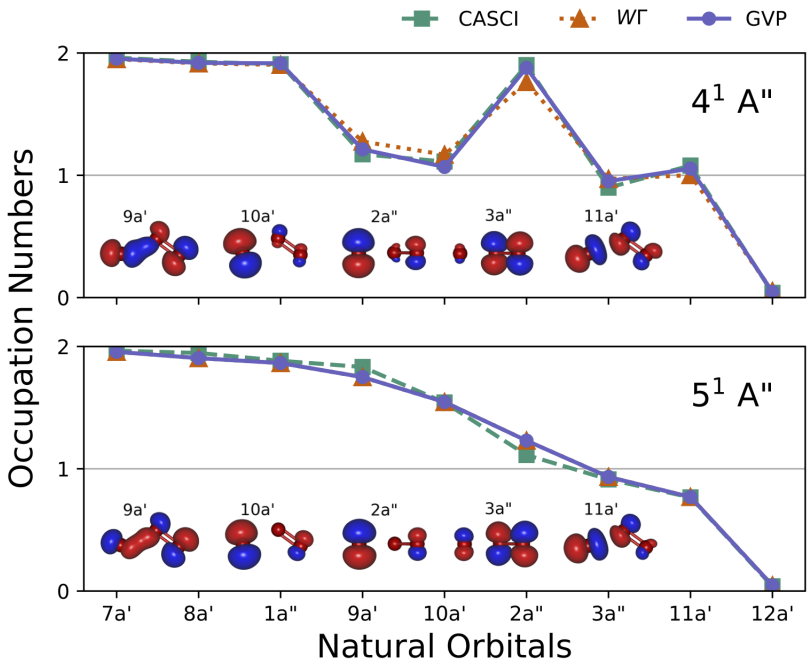


Figure 2.3. Natural orbital occupation numbers for the 4^1A" and 5^1A" excited states of O_3, calculated from the initial CASCI roots and using the $W\Gamma$ and GVP approaches. The insets show the natural orbitals of each state as calculated by the GVP. For each state, $W\Gamma$ converged the orbital energy gradient to 10^{-4} while the GVP converged to 10^{-7}, leading to small discrepancies in the calculated properties.

One property beyond energetics that we can exploit is the fact that different Hamiltonian eigenstates should be orthogonal to each other. When using state-specific optimization and an approximate ansatz, this property will not hold exactly, but should hold approximately. To help find the 5^1A" stationary point, we therefore append an additional component to $\vec{d}$ that (approximately) measures the overlap between the wave function being optimized and the converged GVP 4^1A" state. Our expanded targeting vector in our objective function is now

$$\vec{d} = \left\{ \langle \hat{H} \rangle - \omega, \quad \frac{\vec{b} \cdot \vec{c}}{|\vec{c}|} \right\} \tag{2.17}$$

in which $\vec{c}$ is the CI vector for the wave function being optimized and $\vec{b}$ is the normalized CI vector for the converged 4^1A" stationary point. The new component is only an approximation to the wave function overlap, of course, as it does not account for differences in the shapes of the molecular orbitals in the two wave functions. However, we do not need it to be exact. We only need it to be good enough to push the optimization into the basin of convergence for the

5^1A" stationary point, so that when μ goes to zero in the final stage of GVP optimization, correct convergence is achieved.

Using the expanded targeting vector from Eq. (2.17) led to a successful GVP optimization in which we again started from the 4th ^1A" CASCI root, but this time converged successfully to an energy stationary point for the 5^1A" state. As seen from the overlap data in Table 2.2, while the primary excitation character is easily assignable to the 4th CASCI root, mathematically this stationary point is essentially an equal superposition of the 4th and 5th CASCI roots, revealing that the states remix strongly during state-specific orbital relaxation and that the 4th CASCI root really was a poor initial guess. Near such a crossing of states, small relaxations of the orbital shapes can lead to large changes in the CI coefficients. While the diagonalization procedure of $W\Gamma$ is capable of such changes, GVP is a local search method and thus finds them challenging without the help of additional properties. This motivates more work exploring the abilities of the GVP near energetic crossings and also in seeding it with equal or biased-weighted SA-CASSCF starting points that can start us closer to the solution.

In the end, the two energy stationary points that our GVP finds are made from different mixtures of the 4th and 5th CASCI roots, although with somewhat relaxed orbitals. These stationary points are substantially different from each other but not entirely orthogonal: their exact NOCI-style overlap with each other is 0.3, which is not huge but is not zero either. Thus, although the GVP was successfully able to find SS-CASSCF stationary points for both states in this difficult case, the fact that the final stationary points are not as strongly orthogonal as we might like suggests that the chosen active space could do with enlargement, or at least that a NOCI re-diagonalization of these stationary points may be worthwhile.

2.4.3 MgO

As our third and final example, we use the GVP to find SS-CASSCF energy stationary points corresponding to each of the eight lowest $^1A_1$ CASCI roots in MgO at a bond length of 1.8 Å and with an (8o, 8e) active space. The excited states in MgO present a challenging array of multi-reference and charge transfer character, [168, 169, 170] as can be seen from an inspection of Table 2.3 and Figures 2.5 and 2.6. Some states exhibit both behaviors at once, such as the CT2 state, which is a doubly-excited, double-charge-transfer state in which the most prominent electron configuration accounts for less than half the wave function. SS-CASSCF is an especially appropriate theory in this setting, being able to deal with both the strong post-CT orbital relaxation and the multi-reference character that so often comes along with double excitations. Previous work with state-averaged CASSCF has investigated the lowest excited state in MgO, [171] and in principle dynamic weighting [18] may be able to help in making predictions about the others, but the mix of neutral and ionic character in these states makes standard state averaging hard to recommend, and if one wishes to take dynamic weighting to its limit, one is really asking for SS-CASSCF. However, even when SS-CASSCF is the goal, the method of optimization matters a great deal, with a previous

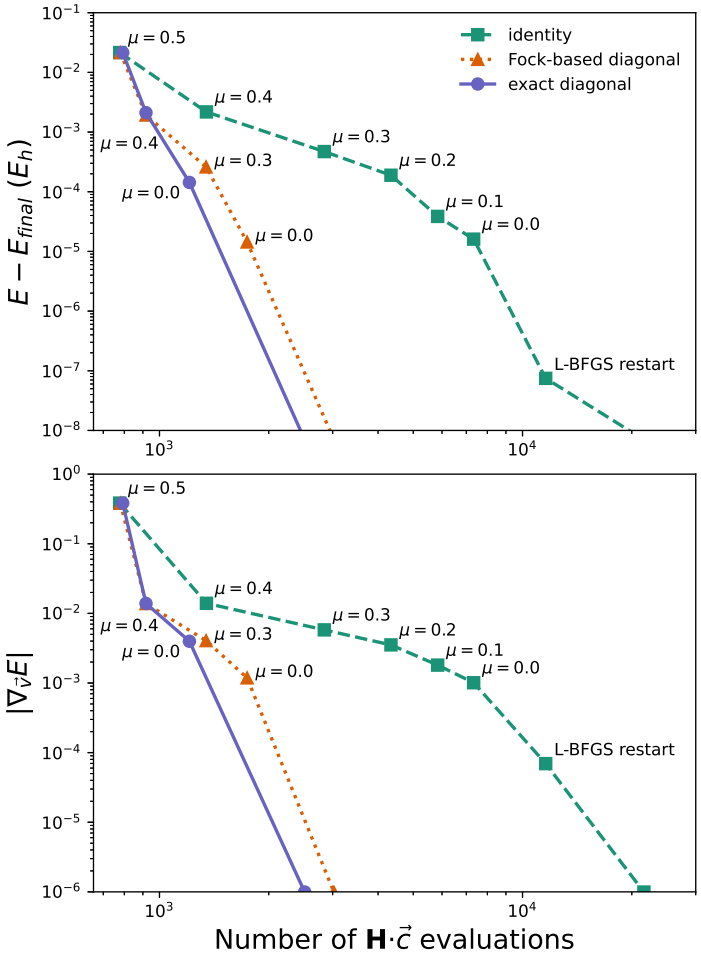


Figure 2.4. Convergence in terms of energy (top) and energy gradient with respect to the variational parameters (bottom) vs the number of Hamiltonian-CI vector contractions for GVP optimizations of the V1 state of MgO. Convergence when L-BFGS starts with an approximate Hessian guess built from the exact diagonal (solid purple line) or Fock-based approximate diagonal (dotted orange line) energy Hessian, is compared to convergence when the identity is used instead (green dashed line). Starting points for new macro-iterations are labeled. The step down in value of μ differs between the GVP variations, as determined by the criteria described in Section 2.3.3. For all optimizations, the first macro-iteration (not shown) uses the identity, $\mu = 0.5$, and freezes the CI parameters to provide some initial orbital relaxation.

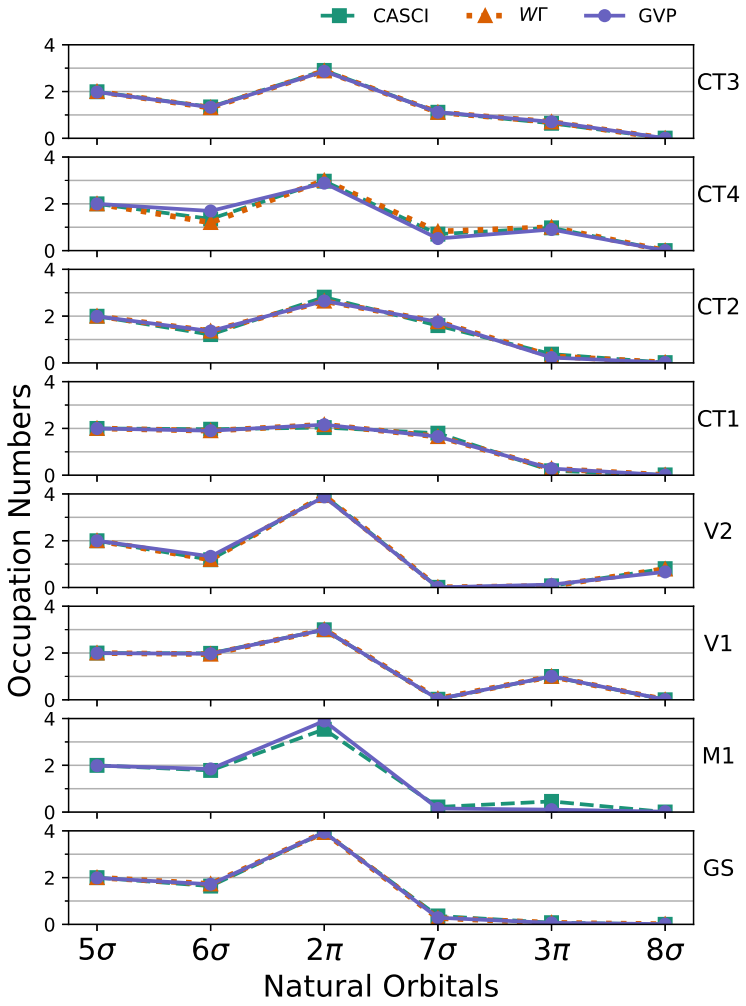


Figure 2.5. Natural orbital occupation numbers for the first eight $^1A_1$ states in MgO, optimized starting from a CASCI-LDA guess with both the $W\Gamma$ and GVP approaches. From bottom to top, the states are displayed in ascending order of the CASCI-LDA energies, although note that due to orbital relaxation, this ordering is not maintained by SS-CASSCF. Note that for both the 2π and the 3π labels, there are two symmetry-equivalent spatial orbitals (i.e. π_x and π_y) and we have grouped them such that for these labels the natural orbital occupations range from 0 to 4.

Table 2.3. Wavefunction data for $^1A_1$ states in MgO, listed from top to bottom in ascending order of the CASCI-LDA energies. Labels (GS, M1, etc) are taken from a previous study. [16] The data include the CASCI-LDA dipole moments μ, wavefunction weight percentages on major components in the LDA orbital basis (the sum of squared determinant coefficients for all determinants of the indicated character), the exact NOCI-style overlaps between the SS-CASSCF stationary points and the initial CASCI-LDA wavefunctions, and the predicted excitation energies.

				Wavefunction Weight %			Overlap		Excitation E (eV)		
State	Label	μ (D)	Primary Excitations	CASCI	$W\Gamma$	GVP	$W\Gamma$	GVP	CASCI	$W\Gamma$	GVP
1^1A_1	GS	-3.95	Aufbau	76.5	81.9	81.9	0.95	0.95	0	0	0
			$6\sigma^2 \rightarrow 7\sigma^2$	12.1	10.9	10.9					
2^1A_1	M1	-5.39	$6\sigma \rightarrow 7\sigma$	41.8	–	56.1	–	0.80	2.48	–	3.11
			$2\pi \rightarrow 3\pi$	25.2	–	1.0					
			$6\sigma^2 \rightarrow 7\sigma^2$	15.1	–	38.4					
3^1A_1	V1	-4.88	$2\pi \rightarrow 3\pi$	68.4	72.5	72.2	0.98	0.98	3.70	4.88	4.88
			$6\sigma, 2\pi \rightarrow 7\sigma, 3\pi$	22.3	19.8	20.0					
4^1A_1	V2	-5.93	$6\sigma \rightarrow 8\sigma$	70.5	44.1	59.9	0.35	0.96	6.46	6.60	8.25
			$6\sigma^2 \rightarrow 7\sigma, 8\sigma$	14.8	22.8	15.0					
			$6\sigma, 2\pi \rightarrow 3\pi, 8\sigma$	5.3	3.8	4.4					
			$2\pi \rightarrow 3\pi$	3.9	0.7	9.1					
			$6\sigma, 2\pi \rightarrow 7\sigma, 3\pi$	2.1	0.6	5.2					
			Aufbau	0.4	17.7	1.0					
			$6\sigma^2 \rightarrow 7\sigma^2$	0.4	6.0	0.5					
5^1A_1	CT1	3.84	$2\pi^2 \rightarrow 7\sigma^2$	62.8	60.7	60.5	0.92	0.92	7.15	6.57	6.57
			$2\pi^2 \rightarrow 7\sigma, 8\sigma$	13.3	7.4	7.5					
			$2\pi^3 \rightarrow 7\sigma^2, 3\pi$	8.7	7.8	7.8					
6^1A_1	CT2	3.93	$2\pi^2 \rightarrow 7\sigma^2$	30.2	44.0	44.1	0.91	0.91	7.62	7.30	7.30
			$6\sigma^2 \rightarrow 7\sigma^2$	16.9	14.7	14.7					
			$6\sigma, 2\pi \rightarrow 7\sigma, 3\pi$	13.9	8.0	8.0					
7^1A_1	CT4	2.33	$6\sigma, 2\pi \rightarrow 7\sigma, 3\pi$	47.1	70.7	52.6	0.30	0.90	8.07	11.65	8.69
			$6\sigma^2, 2\pi \rightarrow 7\sigma^2, 3\pi$	27.0	13.1	24.5					
8^1A_1	CT3	3.66	$2\pi \rightarrow 3\pi$	19.0	16.0	7.9	0.91	0.88	8.16	8.39	8.54
			$6\sigma, 2\pi \rightarrow 7\sigma, 3\pi$	17.4	23.8	31.6					
			$2\pi^2 \rightarrow 7\sigma^2$	16.6	12.8	17.3					
			$6\sigma \rightarrow 7\sigma$	10.0	4.0	1.6					
			$6\sigma^2, 2\pi \rightarrow 7\sigma^2, 3\pi$	8.6	10.1	9.9					
			$2\pi^2 \rightarrow 3\pi^2$	5.9	8.0	6.0					

Table 2.4. Two representative attempts at achieving SS-CASSCF convergence in MgO's M1 state by SA-CASSCF with shifting weights via Molpro version 2019.2 with default SA-CASSCF optimizer settings (aside from the use of biased SA weights). Each attempt starts with an equal-weight SA-CASSCF (seeded with LDA orbitals) and then, for each additional row in the table, uses the previous SA-CASSCF's result as the guess for a new calculation with more biased weights. A 4-state SA was used to simplify the problem by avoiding the states with CT character, but even with this simplification we were not able to get closer than having about 90% of the weight on the target state before root flipping prevented SA-CASSCF from converging. The converged SS-CASSCF energy for M1 found by GVP is -274.403367 E_h.

Attempt 1				
Energy (E_h)	Weight 0	Weight 1	Weight 2	Weight 3
-274.371506	0.250	0.250	0.250	0.250
-274.376089	0.200	0.400	0.200	0.200
-274.384705	0.100	0.700	0.100	0.100
-274.396991	0.050	0.900	0.050	0.000
no convergence	0.025	0.950	0.025	0.000
Attempt 2				
Energy (E_h)	Weight 0	Weight 1	Weight 2	Weight 3
-274.371506	0.250	0.250	0.250	0.250
-274.378506	0.300	0.400	0.200	0.100
-274.390694	0.300	0.600	0.050	0.050
-274.390525	0.400	0.600	0.000	0.000
-274.392357	0.300	0.700	0.000	0.000
-274.394417	0.200	0.800	0.000	0.000
-274.397643	0.100	0.900	0.000	0.000
no convergence	0.050	0.950	0.000	0.000

study showing that simple root selection fails to converge to the initially targeted state in state-specific optimizations of all seven of the lowest $^1A_1$ excited states. [16] Similarly, we find that a shifting-weight SA-CASSCF approach struggles with root flipping in some of these states, as shown in Table 2.4. Using a careful analysis based on NOCI overlaps, we find that, while the WΓ optimization method is more effective, it still fails to locate an appropriate stationary point for three of these seven excited states. By adding the GVP approach to our toolbox, however, we are able to find good energy stationary points for the ground state and all seven excited states.

Before getting into the state-by state details, let us first emphasize the value of supplying L-BFGS with our approximate diagonal form for the initial objective function Hessian as opposed to the identity matrix. For this comparison, as for all the optimizations in this section, our starting point is a particular root from a CASCI calculation carried out in the

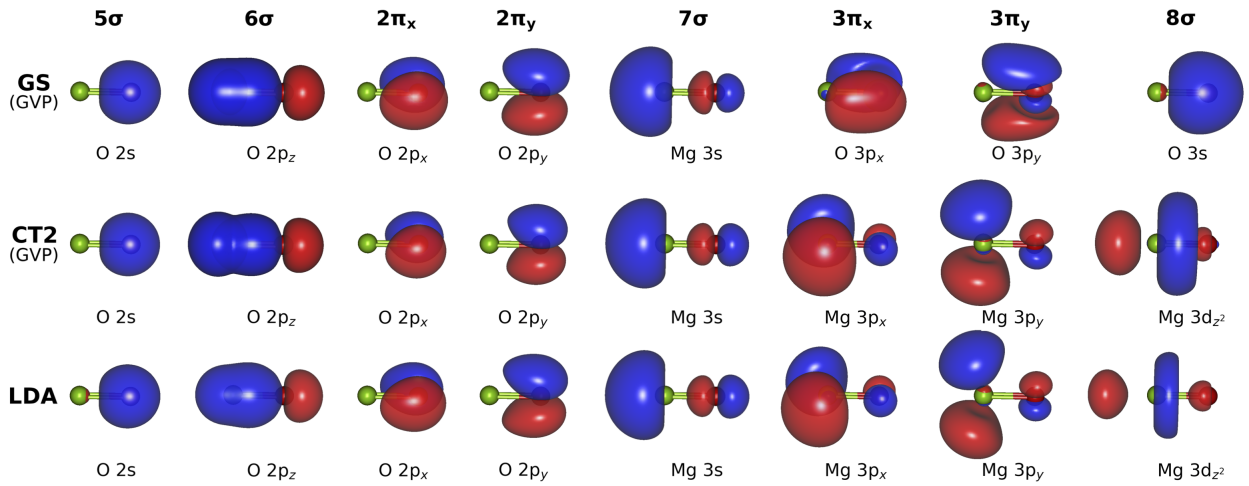


Figure 2.6. The MgO active orbitals in the LDA guess (bottom row) and the SS-CASSCF stationary points for CT2 (middle row) and the ground state (top row). Each image has the Mg atom at left in green and text indicating the orbital's primary character.

LDA orbital basis (denoted as CASCI-LDA), with the active space chosen as the lowest four LDA orbitals of σ character plus the lowest four of π character, as seen in Figure 2.6. These active orbitals can be roughly characterized as the O 2s and 2p and the Mg 3s, off-axis 3p, and $3d_{z^2}$ orbitals. The Mg 1s, 2s, and 2p and the O 1s orbitals are held closed but not frozen. As seen in Figure 2.4, employing either version of our diagonal Hessian approximation speeds up the optimization convergence for the V1 state by more than an order of magnitude relative to using the identity matrix. Similar speed ups were observed for other states as well. There is still room for improvement, however, and so in future it will be interesting to investigate combinations of GVP-based L-BFGS with more standard tools like Davidson CI steps and more traditional orbital optimizations.

Turning now to stability, we find that, with this new GVP optimization method in hand, we can now locate stationary points for all eight of the lowest $^1A_1$ states, as shown in Table 2.3. The ground state is the simplest, and indeed all optimization methods – including GVP, WΓ, SRS, and the default PySCF ground state CASSCF solver – come to the same stationary point. The lowest excited state (M1) is a more significant case, as no previous method has to our knowledge been able to locate the full (orbital + CI) energy stationary point for this state. Despite its careful root tracking approach, WΓ collapses to the ground state when trying to target the M1 state starting from the corresponding CASCI-LDA root. In contrast, GVP has no trouble with this state, finding a stationary point that, based on its NOCI overlap with the starting CASCI-LDA root, clearly corresponds to the excited state being sought. Turning to the V1 and CT2 states, both GVP and WΓ work well, arriving at the same stationary points that, again, have large overlaps with the CASCI-LDA excited states used to initiate the optimizations and define which excited state we are after.

The V2 and CT4 states both represent failures for the WΓ approach, however, which was not obvious in the previous study [16] as a natural orbital occupation analysis (Figure 2.5) makes it appear that the stationary points arrived at are a match for the states being sought. However, NOCI overlaps, which we have now evaluated and which are a more direct measure of wave function similarity, show that in both V2 and CT4, WΓ converges to a stationary point that is of a very different character than the excited state in question. GVP, on the other hand, finds stationary points for these states that have large overlaps with the starting CASCI-LDA roots and so clearly match the states being sought. In CT1, we have our one example in MgO in which the simplest use of the GVP (energy targeting only) fails to find a stationary point, the optimization getting stuck at an energy gradient norm of roughly 10^{-4}. However, WΓ works in this case, and GVP can be improved either by expanding the vector $\vec{d}$, as we did in the upper ozone state, or by improving the initial guess, which is the approach we take here. If we supply slightly better orbitals by taking them from the output of the second macro-iteration of WΓ (but still using the CASCI-LDA CI vector guess so as not to give GVP too much help) we find that the GVP optimization is able to converge to the same stationary point as found by WΓ. The final state we are looking at, CT3, is an even more interesting case, in which WΓ and GVP find two different stationary points, both of which have strong overlap with the sought after state. The difference between these stationary points is in the 8σ orbital, which in the GVP stationary point has O 3s character but in the WΓ stationary point has Mg $3d_{z^2}$ character. Given their large overlaps with the initial CASCI-LDA root and their large overlap of 0.93 with each other, they both appear to be approximations of the same Hamiltonian eigenstate and thus a good example of how nonlinear wave function forms can have more stationary points than there are physical eigenstates. Rather than try to choose between them, we see this as a case that indicates the active space is, at least for this state, at least one orbital too small.

As in other types of CASSCF, multiple solutions can exist when the highest energy active orbitals are only slightly occupied and it is possible to get similarly good wave functions when swapping one or more of them with low-lying virtual orbitals. This issue can cause multiple nearby minima in both ground state and SA-CASSCF, although it is entirely case by case whether swaps between the least occupied active orbitals and the lowest virtual orbitals move the optimization between different local minima or simply move it around within the same basin of convergence surrounding a single minimum. Our results for CT3 provide evidence that something like the multiple-minima issue can occur for excited states in SS-CASSCF, with two very similar stationary points differing by a swap between low-lying virtuals and high-lying active orbitals. In the case of CT3, one might prefer the $3d_{z^2}$ stationary point on the basis that it contains only valence orbitals in its active space, but applying such logic in general is not straightforward. Indeed, all optimization methods we have tried (including the default implementation in PySCF) agree that, after state-specific optimization, the ground state active space displayed in Figure 2.6 contains orbitals with O 3s, $3p_x$, and $3p_y$ character, having swapped them in for the LDA Mg $3p_x$, $3p_y$, and $3d_{z^2}$ valence orbitals that were present in the the initial guess. What is essentially going on here is that, if only a subset of the active orbitals need to have significant occupation in order to

capture the strong correlation effects in a given state, then, for that state, the choice for the remaining active orbitals that will give the lowest energy is whichever ones provide the best ability to capture some weak correlation, and there is no particular reason that these will be valence orbitals. In the ground state, it makes some sense for the O 3-shell orbitals to be more effective for this purpose than the unoccupied Mg valence orbitals, as the ground state concentrates the electrons on the O atom, putting a premium on orbitals that can help describe weak correlation effects in its vicinity. Another well-known example of this issue, although not in play here, is the double d-shell effect,[172, 173] where it is often wise to include non-valence d orbitals in the active space for transition metal compounds ahead of some orbitals that are formally valence orbitals. As in ground states or state averaging cases with multiple minima, the best approach to removing the ambiguity between CT3's two stationary points is probably to expand the active space. By doing so, the orbitals that are competing for inclusion in the active space and leading to multiple stationary points can all be included, at which point we expect the two stationary points would merge into one. From an optimization perspective, this would amount to the two minima on the $|\nabla_{\vec{v}} E|^2$ surface joining into a single minimum with a single basin of convergence. Certainly this must happen in the limit that the active space expands CASSCF into FCI, but we suspect that in this case it will happen immediately upon allowing both the O 3s and Mg $3d_{z^2}$ orbitals to be in the active space simultaneously.

2.5 Conclusion

We have shown that excited-state-specific optimization of the CASSCF ansatz via the minimization of a generalized variational principle allows the desired excited state stationary points to be located and tightly converged in multiple challenging scenarios. The GVP consists of the square norm of the energy gradient along with a steering term that allows approximately known properties of the desired state to guide the optimization to its energy stationary point. The form permits a very broad variety of properties to be employed, and in this study we have used estimates for the energy and, in one particularly challenging case, rough orthogonality against another state for this purpose. By achieving state-specific optimization with the GVP, situations where this approach could be especially helpful include cases where state-averaging is frustrated by root flipping, high-lying states where it is not practical to resolve all lower-lying states, avoided crossings, and states displaying both strongly correlated character and strong orbital relaxations, as in some core, charge transfer and doubly excited states.

In our results, we find that the GVP approach is capable of converging to the correct stationary point in excited states of LiH, ozone, and MgO in which root flipping is present. Its tighter convergence than uncoupled two-step methods produces energies in LiH that are significantly closer to FCI, and its root-targeting capabilities allow it to match the efficacy of the recently developed WΓ method in a nearly degenerate pair of states in ozone. In MgO, it was not previously possible to find the correct stationary points for three excited singlet

states in the symmetric representation of the computational point group. With the addition of the GVP approach, all three of these missing stationary points have been found.

Looking forward, there are a number of promising directions worth pursuing. First, this study limited itself to using quasi-Newton optimization of the GVP objective function, which is illuminating but almost certainly not the most efficient approach given the historical dominance of the Davidson algorithm when dealing with CI coefficients. Methods that combine the flexibility and reliability of GVP minimization with the efficiency of Krylov subspace eigensolvers are thus a priority for future method development. If sticking with a quasi-Newton approach, directions to consider for improving optimization efficiency include correcting the L-BFGS gradient history when shifting the orbital reference throughout the optimization, as well as delving into approximate initial Hessians that retain more of the CI-orbital coupling. Second, CASSCF energetics are rarely quantitative due to a lack of treatment of weak correlation effects. With the GVP approach able to provide excited state stationary points in a wider range of cases than was previously possible, it will be interesting to perform more extensive tests on what benefits this can offer to post-CASSCF weak correlation methods. Whatever these directions uncover, it is becoming increasingly clear that it is possible and often desirable to achieve fully excited-state-specific quantum chemistry in a wide variety of single-reference and multi-reference methods.

2.6 Acknowledgements

This work was supported by the National Science Foundation's CAREER program under Award Number 1848012. Calculations were performed using the Berkeley Research Computing Savio cluster and the Lawrence Berkeley National Lab Lawrencium cluster. R.H. acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1752814 and DGE 2146752. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Chapter 3

CASSCF in thioacrolein: adiabats, diabats, discontinuities, oh my!

3.1 Abstract

We apply fully excited-state-specific complete active space self consistent field theory to the potential energy surfaces of excited states in thioacrolein to help elucidate how this approach differs from state-averaging. With both doubly excited and charge transfer character, thioacrolein's low-lying excitations are an excellent testing ground for multi-reference methods that can fully relax orbital shapes for individual excited states. Near an avoided crossing, we find that different computational protocols can deliver adiabatic or diabatic surfaces. In unconstrained excited state geometry optimization, we find qualitative differences between the state-specific and state-averaged approaches in the second singlet excited state, whose primary components are double excitations that transfer charge away from the sulfur atom.

3.2 Introduction

Predicting the shapes of electronically excited states' potential energy surfaces is a central way in which quantum chemistry contributes to the study of photochemistry. Whether one is focused on Frank-Condon analysis, [174] photoisomerization pathways, [175] nonadiabatic couplings, [176] or the locations of conical intersections, [177] understanding an excited state's potential energy surface (PES) is crucial in fields ranging from astrochemistry [178] to biomedicine [179] to catalysis. [180] For example, the photoisomerization of rhodopsin chromophores is responsible for human vision and is relevant for nanotechnology applications like molecular switches. [181, 182] Method development in quantum chemistry has often been guided by the need for more accurate excited state PESs, but, despite substantial progress, the most effective and widely used excited state methods still bear major shortcomings in this area.

Linear response methods, especially those that build up from a single-reference ground state picture, often struggle when a photochemical process leads to molecular geometries

where their ground state starting points break down. For example, time dependent density functional theory [94, 183] (TD-DFT) has been widely successful in many applications, but can produce qualitatively incorrect PES shapes near conical intersections. [115] This difficulty is not entirely surprising, given that the ground state is often strongly multi-reference at such geometries and so not amenable to treatment with the single-determinant framework of ground state Kohn-Sham theory. Although it has some advantages over TD-DFT, equation of motion coupled cluster theory [87] faces a similar formal dilemma, as it may inherit problems from its ground state coupled cluster starting point at bond-breaking geometries (e.g. during ring openings) where its Hartree Fock reference is qualitatively incorrect. Although these theories can be modified, for example by spin-flip methods, [184] to alleviate these concerns in some settings, these difficulties have long driven interest in multi-reference theories that can directly tackle situations in which the ground and/or excited state involve the strong mixing of multiple electronic configurations.

Complete active space self consistent field theory [48, 49, 50, 51] (CASSCF) is perhaps the most prominent example of multi-reference approaches to modeling excited states and their PESs. It has been widely deployed in the study of photochemistry, with example uses ranging from conical intersections [177, 115] and ring openings [123, 185, 186] to quantum dynamics [187, 188] and molecular switches. [189] With its explicit inclusion of all electronic configurations within an active space of the most relevant orbitals, CASSCF is able to produce qualitatively correct results in multi-reference settings where TD-DFT and equation of motion coupled cluster cannot. However, CASSCF also has its own frustrations, some of which are obvious and well known (choosing a "good" active space is not always straightforward) and others of which are more subtle. On the more subtle side, CASSCF can suffer from non-analytic features in the PES and, especially in excited states, from convergence difficulties. The latter are often driven by root flipping, [190, 171, 191] in which overall convergence is prevented by configuration interaction (CI) roots reordering after an orbital optimization step. Non-analytic features, on the other hand, can occur in either ground or excited states when, during a smooth variation of the molecular geometry, a strongly occupied (unoccupied) active orbital discontinuously swaps with a closed (virtual) orbital. Such swaps, driven by one set of orbitals newly edging out another in the competition to capture the most correlation, typically do not create a qualitative change in the state, but they nonetheless can create kinks in the PES since the swap in orbitals is abrupt.

The most widely used approach to excited states in CASSCF is state-averaging (SA). [111] In this approach, one adjusts the orbital shapes to minimize the (possibly weighted) average energy of a specified set of CI roots, for example the lowest four. Doing so offers a number of advantages: the states in question are all described in the same orthonormal orbital basis and so matrix elements are relatively straightforward, root flipping between states with equal weights no longer creates convergence problems as it does not affect the average energy, and the same types of highly efficient coupled two-step optimization methods [17, 84] that work so well in ground states can be employed. However, SA also has its drawbacks. Root flipping can still create convergence issues when non-equal weights are employed, and can even lead to discontinuities on PESs when the top state in the average

flips with the next state up during a molecular geometry change. [8, 9] The use of one set of orbitals, although convenient, can also create imbalanced treatments if the states included in the average benefit from very different orbital relaxations, as is the case for local versus charge (CT) transfer states. If there are four local states and one CT state in the average, that CT state is likely not seeing the repolarization of the orbitals that one would expect from the creation of hole and particle. This biasing can have serious consequences in the PES, as occurs in 4-aminobenzonitrile (ABN), where SA-CASSCF predicts a qualitatively incorrect relaxed geometry for the low-lying CT state. [120]

Multiple approaches have been developed to address SA-CASSCF's shortcomings. Some attempt to remove orbital-optimization issues, such as the PES discontinuities born of root flipping, by replacing CASSCF with a complete active space CI (CASCI) run in a pre-optimized basis of Hartree-Fock or DFT orbitals. [192] However, this approach inherits issues from DFT and Hartree-Fock, which often have discontinuities of their own in multi-reference settings. [10] Others allow for larger active spaces, [24, 55, 27, 20] which in principle can be expected to reduce issues with orbital swapping. Still other approaches work to make excited state CASSCF more state-specific. For example, dynamically weighted SA adjusts the weights on the fly to focus primarily on the states that drive the photophysics, such as those involved in an avoided crossing, which helps those states get the orbital relaxations they need while maintaining a multi-state approach. [18, 8, 9]

Recently, this push towards more excited-state-specific CASSCF has led to the development of methods that optimize the orbitals for an individual excited state, fully eschewing averaging between states. These approaches can be put in two rough categories: those that separate CI and orbital updates in a traditional two-step optimization procedure and those that fully couple the optimization of both. Like the selection of occupied orbitals in the maximum overlap method for Δ-SCF, [66, 67] the two-step methods must, after each orbital update, decide which CI root is the "right" one to use for the next orbital update. The WΓ method [16, 121, 193] employs an energy- and density-matrix-based criteria for this root selection, which has proven effective in many but not all root flipping scenarios. Fully coupled methods, such as the generalized variational principle (GVP) approach, [21, 194] do not face the two-step issue of needing to select a root at each iteration as they simply perform quasi-Newton updates of the orbitals and CI parameters together. As such, they are even less prone root flipping, but, as studied carefully by Burton, [33] may produce diabatic instead of adiabatic solutions due to their optimization algorithm's lack of re-diagonalization. Aside from the ability to fully relax orbitals for the state in question (which has proven crucial in some CT settings [121, 193]), another clear advantage that fully state-specific (SS) CASSCF methods offer is that, since they produce energy stationary points, their nuclear energy gradients are much more straightforward than those of SA-CASSCF. [121]

To further elucidate the differences in behavior between SA-CASSCF and SS-CASSCF and to further test adiabatic vs diabatic behavior in the latter, this study explores the PESs of the low-lying states of thioacrolein, a biologically active [195, 196] derivative of ethylene that offers a rich tableau of π-bond twisting, CT character, and avoided crossings. Key questions include which excited states adopt a bent versus planar structure, which

approaches to SS-CASSCF produce diabats versus adiabats, and which approaches show or don't show non-analytic features on the PES. We find that the SA and SS approaches can produce qualitatively different PES features, that non-equal-weight SA is especially prone to discontinuities, and that a mixture of SS and SA methodologies can increase the likelihood of producing smooth, adiabatic SS-CASSCF surfaces.

3.3 Theory

3.3.1 The CASSCF Ansatz

To treat multi-configurational systems, CASSCF separates the molecular orbitals into three categories (closed, active, and virtual) and then constructs a linear combination of the Slater determinants formed by doubly occupying the closed orbitals and then arranging the remaining electrons in all possible ways within the active orbitals. The coefficients of this linear combination, as well as the shapes of the molecular orbitals, are then varied to optimize an objective function, such as the ground state energy or, as in SA, the average energy of the first so many CI roots. Mathematically, this ansatz can be written as

$$|\Psi\rangle = e^{\hat{X}} \sum_i c_i |\phi_i\rangle \tag{3.1}$$

$$\hat{X} = \sum_{pq} X_{pq} \hat{a}_p^\dagger \hat{a}_q \tag{3.2}$$

where c_i are the CI coefficients, $|\phi_i\rangle$ are the Slater determinants, and $\exp(\hat{X})$ is an orbital rotation operator [29, 21] defined by the anti-symmetric matrix $\boldsymbol{X}$. As core-core, active-active, and virtual-virtual orbital rotations do not change the variational flexibility of the ansatz (and thus have no effect on the CASSCF energy), their blocks of $\boldsymbol{X}$ are neglected, leading to the parameterization $\vec{x}$ shown in Figure 3.1. Altogether, the CASSCF wave function's variational parameters are the set $\vec{\nu} = \{\vec{c}, \vec{x}\}$.

3.3.2 CASSCF wave function optimization

In both ground state and SA CASSCF, the variational parameter sets $\vec{x}$ and $\vec{c}$ are typically updated separately during optimization via an iterative two-step procedure that minimizes either the energy of one state or the average energy of multiple states. [16]

$$E_{\text{SA}} = \sum_{I \in \text{SA}} w_I \langle \Psi_I | \hat{H} | \Psi_I \rangle \tag{3.3}$$

One of the "steps" is typically the Davidson algorithm [41] for CI parameters, while the other is an augmented Hessian update for the orbitals that approximately accounts for some orbital-CI coupling effects. [17] A notable challenge for two-step methods is that, following the orbital update, the ordering of the CI roots may change compared to the previous

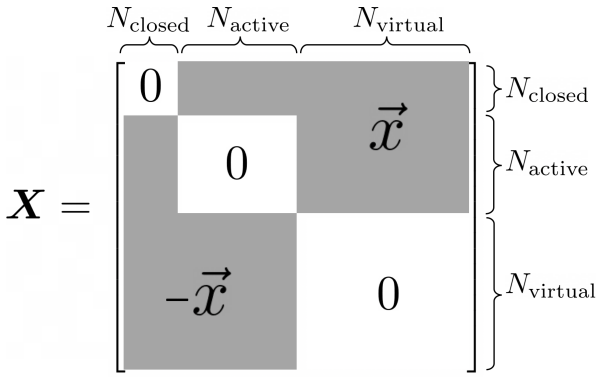


Figure 3.1. Orbital rotation coefficient matrix $\boldsymbol{X}$ where the solid shaded areas represent nonzero variational parameters or their negative transpose.

iteration, an occurrence that is commonly referred to as root flipping. In some molecules, especially when using unequal weights w_I in SA or going fully SS with all the weight on one excited state, root flipping can prevent the two-step method from converging. SA can counter this problem by using equal weights, assuming the root flipping is between states included in the average. In fully SS approaches, one can instead attempt to track the root through the reordering, for example using maximum-overlap methods or the energy and density matrix.[16] However, even these root tracking methods can fail when two roots become nearly degenerate and mix strongly, and equal-weight SA does not protect against convergence failures born of flipping between the highest root in the average and the lowest root not in the average. These challenges, as well as the potential benefits of fully SS orbital optimization, have motivated work on optimization methods that do not rely on a two-step approach and instead update orbital and CI parameters together via local-search nonlinear minimization algorithms.

One such approach is to apply quasi-Newton optimization to an excited-state-focused GVP. [194] On the assumption that an approximate but high-quality ansatz will have energy stationary points corresponding to the true excited states, [105, 65, 66, 67] the GVP approach constructs a shifting objective function that mixes minimization of the energy gradient norm with other properties that act to guide minimization towards the desired stationary state.

$$L = \mu|\vec{d}|^2 + (1-\mu)\left|\nabla_{\vec{v}} E\right|^2 \tag{3.4}$$

In general, $\vec{d}$ is a vector of differences between expected properties of the desired state and the actual value of these properties for the current wave function, allowing everything from energy to dipole to overlaps with other states to be used to guide the optimization. [21] In this study, we keep things relatively simple with the one-element vector $\vec{d} = \{\omega - E\}$ that measures the difference between the current energy and a guess ω for the energy of the stationary point we are seeking. The parameter μ, which may start at a value like 1/2,

is gradually reduced to zero over the course of the optimization, allowing the guiding $|\vec{d}|^2$ term to help move us towards a specific stationary point early in the optimization while still allowing precise convergence to the $\nabla_{\vec{\nu}} E = 0$ stationary point at the end. Recalling that $\vec{\nu}$ contains both the orbital and CI parameters,

$$|\nabla_{\vec{\nu}} E|^2 = \sum_i \left|\frac{\partial E}{\partial c_i}\right|^2 + \sum_j \left|\frac{\partial E}{\partial x_j}\right|^2 \tag{3.5}$$

we see that a quasi-Newton minimization of L (we use limited-memory Broyden-Fletcher-Goldfarb-Shanno L-BFGS [154, 155, 156, 157, 82] with an approximate diagonal hessian [194]

$$\frac{\partial^2 L_\mu}{\partial v_i^2} \approx 2\mu \left[(E-\omega)\mathcal{H}_{ii} + \left|\frac{\partial E}{\partial v_i}\right|^2\right] + 2(1-\mu)\mathcal{H}_{ii}^2 \tag{3.6}$$

built using the exact diagonal of the energy Hessian $\mathcal{H}$ and employ a finite-difference technique [73] for gradients of L) will update all the parameters together and will never be faced with the type of root selection dilemma that can arise in two-step approaches. That said, it is important to stress that, as a nonlinear ansatz, CASSCF may have extra, spurious stationary points. [33] Thus, even without root flipping, care may be necessary to ensure an appropriate stationary point is found, an issue we will return to below when discussing how we might aim to find stationary points that behave in a diabatic or adiabatic fashion near avoided crossings.

3.3.3 Nuclear gradients

One advantage of fully SS excited state CASSCF is that, like its ground state counterpart, [197] it produces a wave function whose energy is stationary with respect to all of its variational parameters. As such, the Hellmann-Feynman theorem guarantees that energy gradients with respect to changes in the molecule (e.g. the nuclear positions $\boldsymbol{R}$) or its environment can be evaluated straightforwardly via derivatives of the Hamiltonian.[198]

$$\frac{dE}{d\boldsymbol{R}} = \langle\Psi|\frac{d\hat{H}}{d\boldsymbol{R}}|\Psi\rangle \tag{3.7}$$

In contrast, the energies of the individual states from a SA calculation are not stationary with respect to the orbital parameters, as it is the average energy rather than the energy of an individual state that is minimized. The result is that SA nuclear gradients involve additional terms and the solution of coupled-perturbed CASSCF equations to evaluate the orbital response. While these have been implemented in multiple codes, they substantially increase the effort required to evaluate nuclear gradients when compared to the ground state case. [199, 200]

3.3.4 GVP single point optimization procedure

In detail, the overall GVP-based optimization procedure used in this study is as follows.

1. An initial molecular orbital (MO) basis is chosen, and the necessary electron repulsion integrals in that basis are evaluated via a Cholesky decomposition with a screening tolerance of 10^{-5} Hartrees. [201, 202, 203] Cholesky with this same tolerance is used for all other MO basis integral evaluations needed throughout the optimization.

2. Initial CI coefficients are either taken from a previous calculation or, if we are using equally weighted SA-CASSCF to get the initial orbitals, from its final CI vector for the state in question. The value for ω is set to either the energy of the previous calculation or the SA-CASSCF energy of the targeted state.

3. A series of L-BFGS minimizations (macro-iterations) are performed on L. In each minimization, the L-BFGS history is reset and the initial approximate Hessian is set following Eq. (3.6) using the diagonal of the exact energy Hessian. [194]

    - In the first macro-iteration we set μ to 0.5, hold the CI parameters fixed, and take L-BFGS steps until the orbital gradient satisfies $|\nabla_{\vec{x}} L| < 10^{-2}$.
    - In each macro-iteration following the first, μ is decreased by 0.1, the gradient threshold is tightened by a factor of 10 until reaching its final value of 10^{-6}, and both the orbital and CI parameters are optimized together.
    - If, at the end of a macro-iteration, the largest gradient element is smaller than the current gradient threshold, we immediately set μ to 0 and perform a final minimization with convergence thresholds $|\nabla_{\vec{v}} |\nabla_{\vec{v}} E|^2| < 10^{-6}$, $|\nabla_{\vec{c}} E| < 10^{-5}$, and $|\nabla_{\vec{x}} E| < 10^{-5}$. Otherwise, this final minimization, with these thresholds, is performed when μ is stepped down to 0 in the sixth macro-iteration.

3.3.5 GVP geometry optimization and interpolation

For geometry optimizations and interpolations, in which a guess from a previous geometry is typically available, we test two methods that use the GVP to arrive at SS-CASSCF solutions. Although they are not guaranteed to do so, testing reveals that one of these methods tends to generate diabatic solutions while the other tends to generate adiabatic solutions, and so we will designate them as Method D and Method A respectively. For both methods, we perform geometry optimizations utilizing PySCF modules[161] in conjunction with geomeTRIC[204] for unconstrained optimizations using a trust radius based step control algorithm and convergence criteria (in a.u.) set to $\Delta E < 10^{-4}$, gradient RMS $< 3 \times 10^{-4}$, gradient maximum $< 4.5 \times 10^{-4}$, displacement RMS $< 1.2 \times 10^{-3}$, and displacement maximum $< 1.8 \times 10^{-3}$. Step by step, the two methods are mostly the same, differing only in step 3 as shown here.

1. At the initial geometry, an equally weighted four-state SA-CASSCF calculation is performed, and the targeted state's MO coefficients, CI vector, and energy are stored. We then repeatedly apply steps 2 through 5.

2. If the stored MO coefficients are not already orthonormal at the current geometry, we apply Löwdin orthonormalization to make them so.

3. **Method D:** Perform a single-point GVP optimization, as outlined in Section 3.3.4, with the stored MO coefficients and CI vector used as the initial guess and the stored energy used for ω. Store the final MO coefficients, CI vector, and energy.
   Method A: First, perform an equal-weight 4-state SA-CASSCF calculation, using the stored MO coefficients as the initial guess. Then, select the SA-CASSCF root whose CI vector has the largest dot product with the stored CI vector, and use the SA-CASSCF-optimized MO coefficients and that root's CI vector as the inital guess for a GVP single-point optimization, with that root's energy employed for ω. Upon GVP convergence, store the MO coefficients, CI vector, and energy.

4. For geometry optimizations, we exploit the fact that, as an energy stationary point, the nuclear gradients for a SS-CASSCF excited state may be evaluated with the exact same algorithm as the ground state, and so pass the newly stored MO coefficients and CI vector to PySCF's *mcscf* module to generate analytic nuclear gradients. These gradients, as well as the newly stored energy, are passed through PySCF's *geomopt* module to geomeTRIC, whose geometry optimization algorithm updates the molecular geometry. For geometry interpolations, we instead simply update the molecular geometry to the next interpolation point.

5. If the geomeTRIC convergence criteria are not met or we are not done with the interpolation, we return to step 2.

3.4 Results and Discussion

3.4.1 Computational details

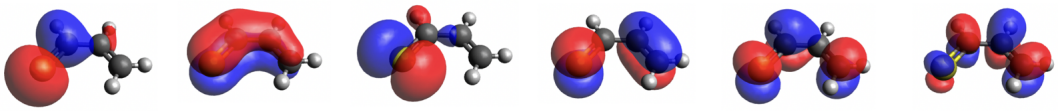
All CASSCF calculations were carried out in the def2-svp atomic orbital basis [205] and employed an (8e, 6o) active space consisting of two lone pair orbitals and the four orbitals in the valence π space. SA-CASSCF results were evaluated with Molpro, [206] while SS-CASSCF used a combination of PySCF, [161] geomeTRIC, [204] and our own in-house GVP code. All molecular orbital analysis and plotting was performed with Avogadro.[207, 208]

3.4.2 Overview of states

In this study we focus on the first three singlet excited states of thioacrolein, which we label as ES1, ES2, and ES3. At the ground state equilibrium geometry (Figure 3.2), which all methods agree is planar, these excited states have $n \rightarrow \pi^*$, $(n, \pi) \rightarrow (\pi^*)^2$, and a mixture

Table 3.1. Wave function characters and active orbitals for the first four states of thioacrolein at the SS-CASSCF ground state equilibrium geometry.

State	Primary Excitations	Dipole Moment (D) SA-25,25,25,25	Dipole Moment (D) GVP	Energy (a.u.) SA-25,25,25,25	Energy (a.u.) GVP
GS	Aufbau	2.1	2.0	-513.2532958	-513.2579257
ES1	$n \rightarrow \pi^*$	0.9	0.8	-513.1796642	-513.1828402
ES2	$n, \pi \rightarrow (\pi^*)^2$	0.9	1.0	-513.0603951	-513.0626141
ES3	$\pi \rightarrow \pi^*$ $\pi^2 \rightarrow (\pi^*)^2$	2.0	2.0	-513.0507352	-513.0529885



n_1 π_1 n_2 π_2 (HOMO) π_3 (LUMO) π_4

of $\pi \rightarrow \pi^*$ and $\pi^2 \rightarrow (\pi^*)^2$ characters, respectively, as shown in Table 3.1. Although the weights of the individual configurations differ slightly between SA and GVP, they broadly agree on the state characters. The ground state is unsurprisingly dominated by the Aufbau configuration, while ES1 is dominated by the $n_1 \rightarrow \pi_3$ transition, giving it strong sulfur-to-C_3H_4 CT character. ES2 shows similarly strong CT character, with roughly the same change in dipole moment relative to the ground state, but is dominated by $(n, \pi) \rightarrow (\pi^*)^2$ double excitations involving the n_2, π_2, and π_3 orbitals, with a minor contribution coming from the $n_2 \rightarrow \pi_4$ single excitation. ES3 exists primarily as a mixture of $\pi^2 \rightarrow (\pi^*)^2$ double excitations and high-lying $\pi \rightarrow \pi^*$ single excitations. Thus, even at this ground state geometry, the low-lying excitations in thioacrolein contain both CT and doubly excited character, making them strong candidates for treatment with methods that can handle both multi-reference character and post-excitation orbital relaxations. Twists away from the planar ground state geometry, to which we will now turn in order to investigate avoided crossing behavior, make the need for multi-reference treatments even more pressing as they introduce strong correlation even into the ground state.

3.4.3 A Small-Twist Avoided Crossing

To investigate how different optimization protocols can favor adiabatic or diabatic character in GVP-based CASSCF, we have focused on a modest twist away from the planar ground state geometry that induces an avoided crossing between ES2 and ES3. Specifically, we have produced a modestly twisted geometry by freezing the $CCCH_{trans}$ dihedral angle at 185 degrees and relaxed the remaining geometric parameters via a biased-weight SA ES2

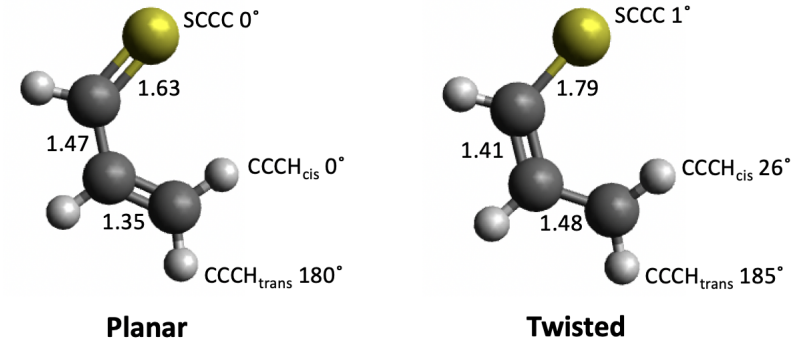


Figure 3.2. Ground state planar geometry (left) and relaxed 5 degree twisted geometry (right). The $CCCH_{trans}$ dihedral was frozen at 185 degrees and the remaining degrees of freedom were optimized using SA-1,1,98,0.

geometry optimization in which the weights are set to 1%, 1%, 98%, and 0% (referred to as SA-1,1,98,0). During this optimization, we see both the C-S and the ethylene C-C bonds elongate and the *cis* H pop out of the C-C-C plane. We have then performed a linear z-matrix interpolation between the ground state geometry and this modestly twisted geometry, both shown in Figure 3.2, along which we find an avoided crossing between ES2 and ES3.

Looking first at the SA-CASSCF results along this interpolation, which are shown in the left-hand panels of Figure 3.3, we make two observations. First, regardless of the SA weighting, SA-CASSCF produces adiabatic potential energy curves for the ES2/ES3 avoided crossing, as is to be expected for a method that explicitly diagonalizes the states against each other. Second, the potential energy curves contain discontinuities for all weightings except for equal weighting (SA-25,25,25,25). These discontinuities are can be explained by root flipping and discontinuous changes in the optimal active space orbital shapes. Looking at the SA-25,50,25,0 curves, the energy jumps when the second and third excited states abruptly swap ordering, changing which gets the 25% and which gets the 0% weighting. In the SA-25,25,50,0 weighting that focuses on the second excited state, the discontinuity is associated with an abrupt swap of the nearly-doubly-occupied n_1 active orbital between a more C-localized shape and a more S-localized shape that coincides with a discontinuous swap in character between ES2 and ES3. This discontinuous change in active space orbital and state character can also be seen in the dipole moments shown in Figure 3.4. A similar abrupt orbital swap also creates a discontinuity in SA-20,20,20,40, although in that case the discontinuity occurs to one side of the avoided crossing and so the crossing itself plays out smoothly, as can be seen in the dipole moments. Interestingly, despite the fact that equal-weight-SA's safety in the face of most root flipping offers no guarantee against the type of abrupt orbital swap seen in SA-20,20,20,40, it nonetheless succeeds in avoiding such difficulties and producing smooth potential energy curves during this interpolation.

Turning now to GVP-based SS-CASSCF, we see in the right panels of Figure 3.3 that, de-

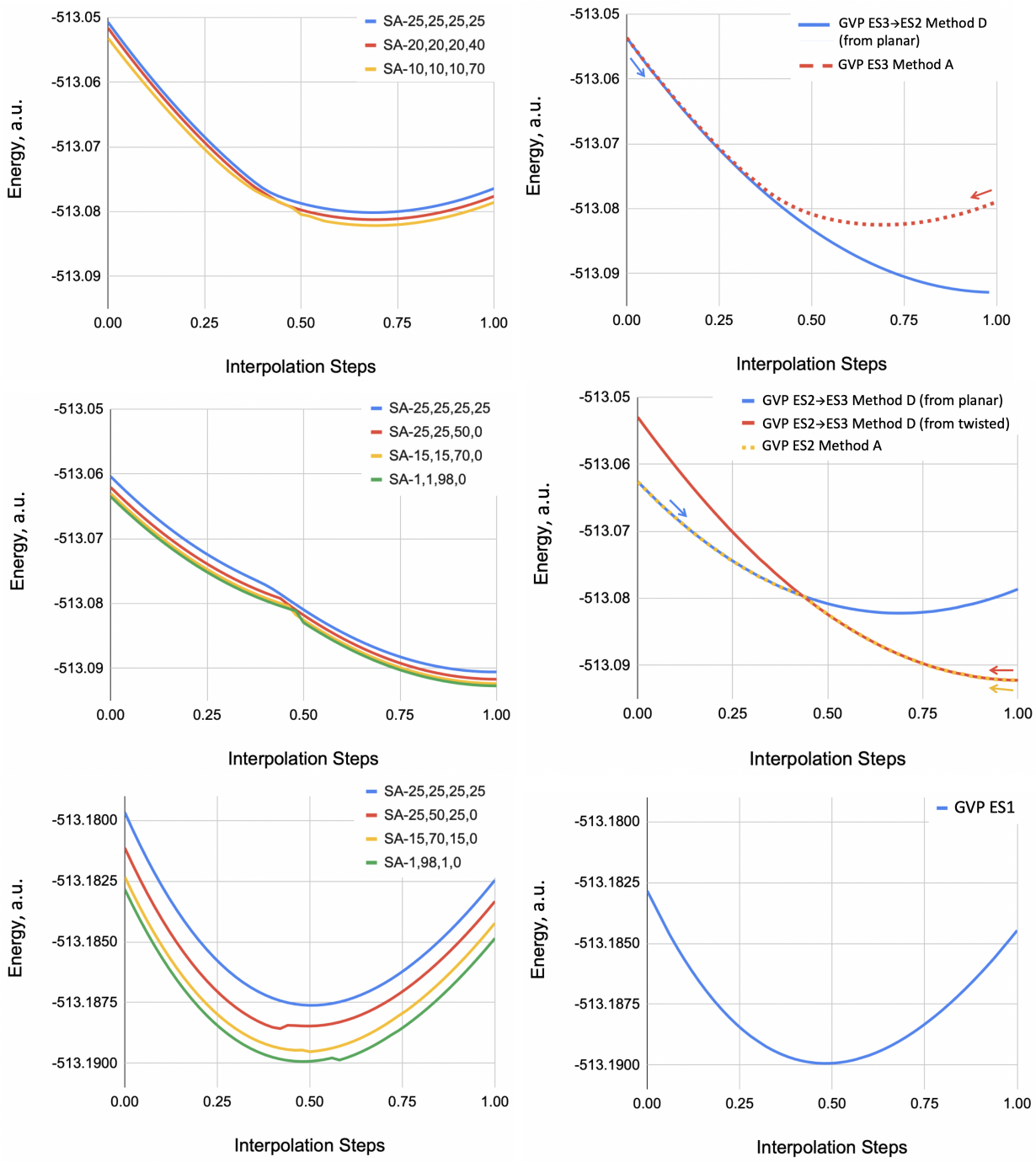


Figure 3.3. Potential energy surface plots of thioacrolein for ES1 (bottom), ES2 (middle) and ES3 (top) across a geometry interpolation between the planar ground state geometry and a relaxed 5 degree twisted geometry. The $CCCH_{trans}$ dihedral was frozen at 185 degrees and the rest of the geometry was optimized using SA-1,1,98,0. The left panel shows the SA results while the right panels shows the GVP.

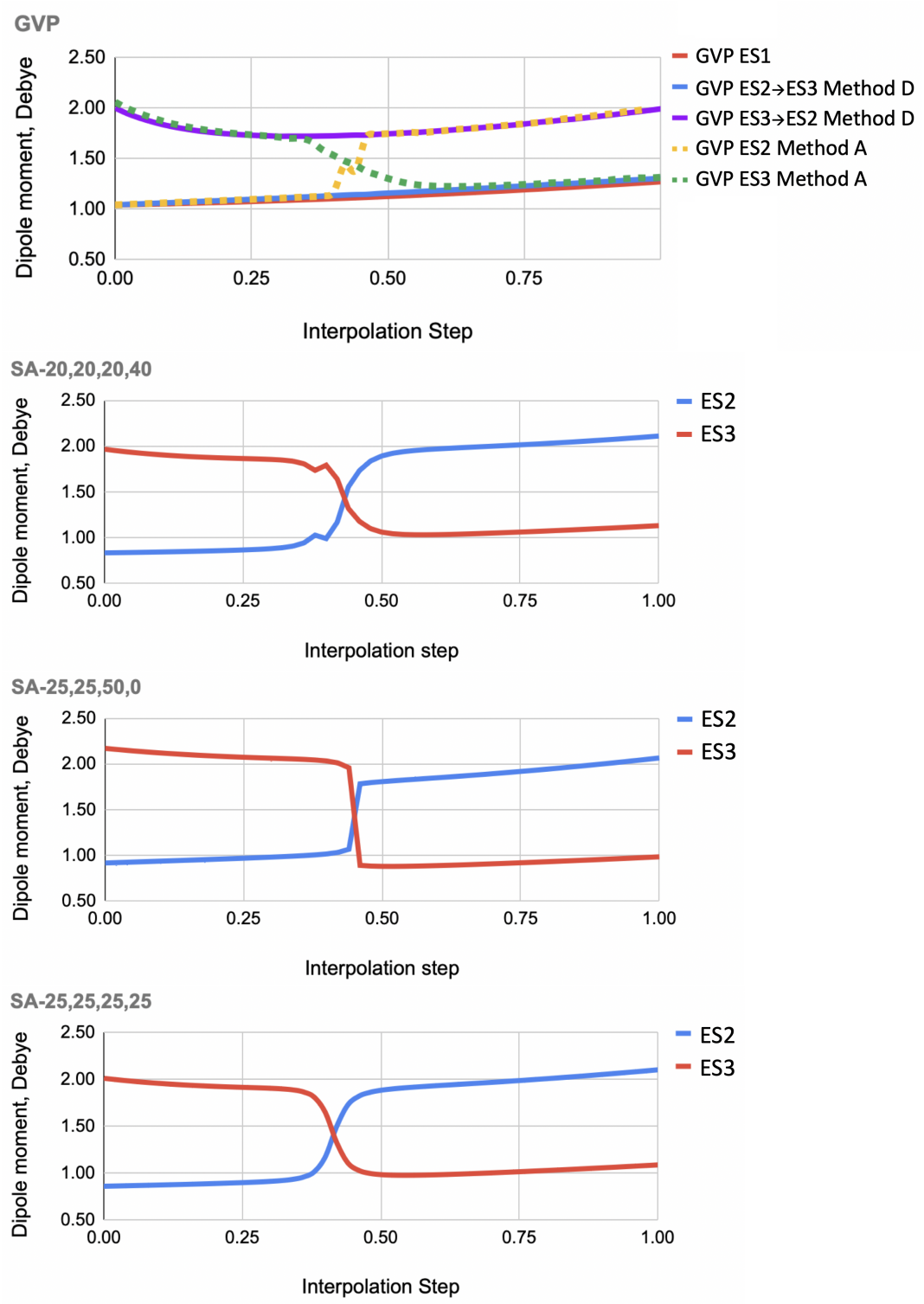


Figure 3.4. Dipole moment across the geometry interpolation between the planar ground state geometry and a relaxed 5 degree twisted geometry. The bottom three plots show the dipole moments of various SA calculations: equal weighting on the bottom, then 50% on the second excited state and then 40% on the third excited state. The top plot shows the dipole moments for the diabats (solid) and the adiabats (dashed) found with the GVP.

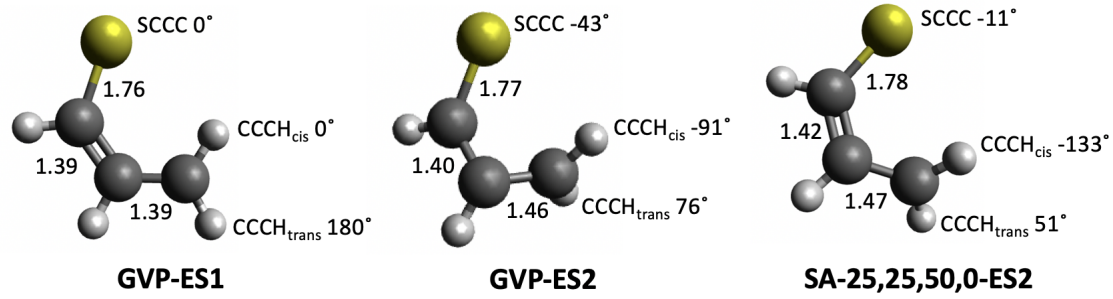


Figure 3.5. Equilibrium geometries for the first and second excited states found using GVP and the most biased successful SA approach.

pending on the computational protocol, one can produce either adiabatic or diabatic curves for ES2 and ES3. Both sets of curves are CASSCF energy stationary points, which adds further evidence to Burton's observations [33] that SS-CASSCF's nonlinear character can cause it to have multiple stationary points associated with an individual excited state. Happily, our computational protocols that one might expect to favor diabatic versus adiabatic results do indeed do so in this case. Method D, which involves no re-diagonalization in its initial guess for a new geometry, produces smooth diabatic curves for ES2 and ES3, as well as a smooth curve for ES1. Although we do not pursue it here, one could turn these into smooth adiabats via a 2x2 non-orthogonal CI rediagonalization. [209] Method A, on the other hand, whose initial guess setup for each new geometry involves a rediagonalization inside a SA-CASSCF calculation, produces SS-CASSCF energy stationary points that are adiabatic in character. The adiabatic versus diabatic nature of the SS-CASSCF results is also emphasized by the dipole moments shown in Figure 3.4. In the Method A results, the adiabatic swap between states with and without $n \rightarrow \pi^*$ components swaps the dipole strengths, because $n \rightarrow \pi^*$ transitions produce CT from the sulfur while $\pi \rightarrow \pi^*$ transitions essentially do not. In the Method D results, the two states maintain a lower and a higher dipole throughout, indicated that they are indeed diabatically maintaining their properties.

Looking ahead to the unconstrained geometry optimizations that we will investigate next, we already see hints of their behavior in this interpolation between planar and modest twisting. Along this interpolation coordinate, both SA-CASSCF and SS-CASSCF agree that ES2 favors a more twisted geometry, and they also both agree that ES1 has its lowest energy not too far from planar. One might predict from this that unconstrained optimization on ES1 would result in a planar geometry while ES2 would be twisted. As we will now discuss, this is exactly what happens in SS-CASSCF, but in SA-CASSCF the picture depends strongly on the chosen weighting.

Table 3.2. Table of geometry optimization data for thioacrolein using various four-state SA weightings and the GVP approach.

Method	Dihedral angle (°) SCCC	$CCCH_{trans}$	$CCCH_{cis}$	S-C bond length (Å)	Dipole moment (D)	$E_{final} - E_{initial}$ (a.u.)
First Excited State						
initial geom.	1.3	-175.0	25.9	1.79	0.89	
SA-25,25,25,25	0.0	-180.0	0.0	1.76	0.99	-0.0103
SA-33,33,33,0	0.0	-180.0	0.0	1.76	0.87	-0.0102
SA-20,60,20,0	0.0	-180.0	0.0	1.76	1.00	-0.0102
SA-1,98,1,0	0.0	-180.0	0.0	1.76	1.18	-0.0103
GVP	0.0	-180.0	0.0	1.76	1.19	-0.0104
Second Excited State						
initial geom.	1.3	-175.0	25.9	1.79	2.06	
SA-25,25,25,25	-3.9	45.9	-141.9	1.78	2.02	-0.0014
SA-33,33,33,0	-5.8	47.2	-139.5	1.78	2.04	-0.0014
SA-25,25,50,0	-11.1	51.1	-132.5	1.78	2.10	-0.0015
SA-23,23,54,0	-45.3	77.8	-91.0	1.77	2.31	-0.0047
SA-20,20,60,0	-44.8	77.5	-91.0	1.77	2.30	-0.0048
SA-15,15,70,0	-44.2	77.0	-91.1	1.77	2.29	-0.0054
GVP	-43.1	76.3	-91.1	1.77	2.25	-0.0068
SA-25,25,25,25 from GVP equil.	-46.7	78.9	-90.4	1.78	2.27	-0.0034

3.4.4 Unconstrained geometry optimizations

Starting from the modestly twisted geometry used as one end of our avoided crossing interpolations, we have performed full, unconstrained geometry optimizations on the ES1 and ES2 PESs using both SA-CASSCF and GVP-based SS-CASSCF. Looking first at ES1, we see in Figures 3.2 and 3.5 that both approaches predict planar geometries with elongated C-S bond lengths. Notably, both Method A and Method D produced the same geometry for SS-CASSCF. Compared to the SS results, the dipole moments of the different SA results vary some, presumably because SA prevents full relaxation of the closed-shell orbitals in the presence of this state's CT. That said, they do appear to limit towards the SS dipole as the weightings are made more and more biased in favor of ES1 as one would expect.

For ES2, the picture is much less straightforward, with Figures 3.2 and 3.5 revealing that, while SS-CASSCF prefers a twisted SCCC dihedral, SA-CASSCF's preferences are a strong function of the weighting. More equal weightings produce SCCC dihedrals close to planar, while more biased weightings produce twists close to those of SS-CASSCF. To make some sense as to what is going on, we first checked that both Method A and Method D give the

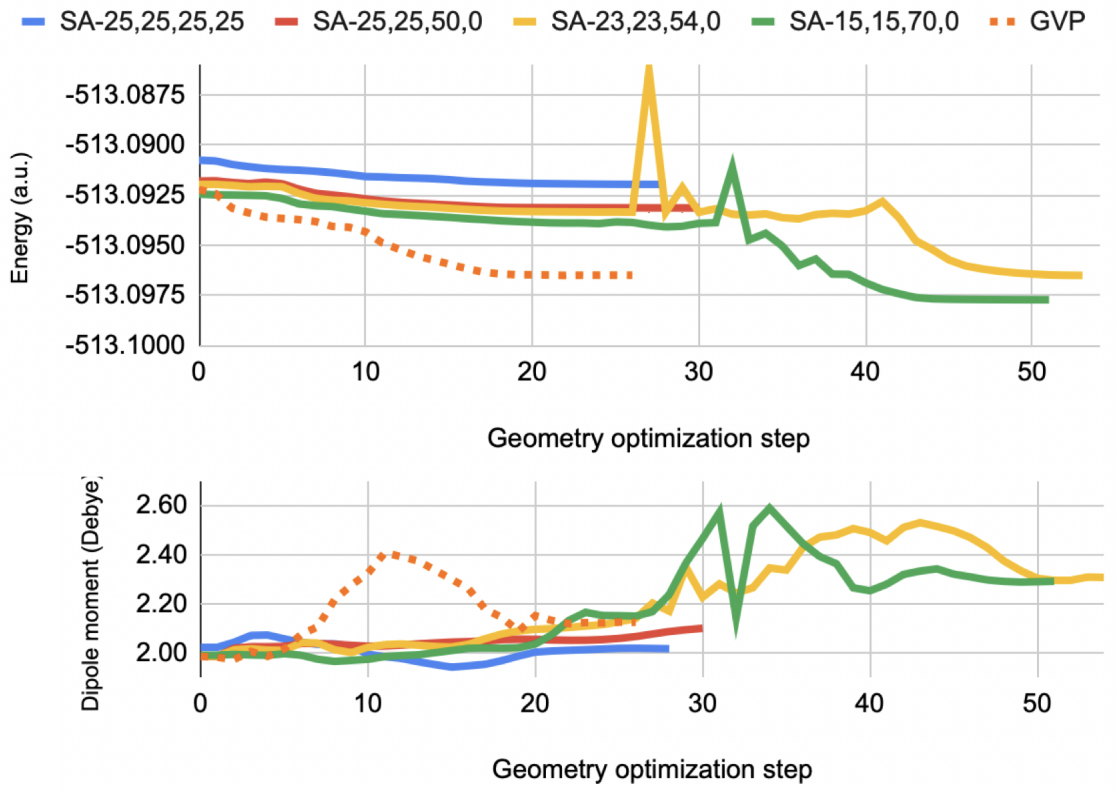


Figure 3.6. Energies and dipole moment throughout the geometry optimizations using the GVP (dashed line) and various biased weighting SA (solid lines) targeting the second excited state.

same geometry, which indeed they do. We then looked at the potential energy curve produced by equal-weight SA-CASSCF when it is evaluated at the same set of geometries that are stepped to during our SS-CASSCF optimization, and, as shown in Figure 3.7, found that the more planar SA-CASSCF geometry is in fact a local minimum. Indeed, if one starts an equal-weight SA-CASSCF optimization from the relaxed SS-CASSCF geometry, it is then able to find the lower-energy, more twisted structure. While more biased-weightings in SA-CASSCf produced structures closer to the GVP structure, a closer inspection of the optimizations reveals that their PESs also differ qualitatively from SS-CASSCF. In particular, biased-weight SA with more than 50% of the weight on ES2 displayed PES discontinuities midway through their geometry optimizations, as seen in Figure 3.6. As we saw in the avoided crossing interpolation, these discontinuities are related to discontinuous changes in the n_1 orbital between more sulfur-based and more carbon-based shapes. SS-CASSCF also sees significant changes in this orbital's shape during its geometry optimization, but they occur smoothly rather than discontinuously.

The situation in thioacrolein's ES2 state thus appears to be similar to that in ABN's CT

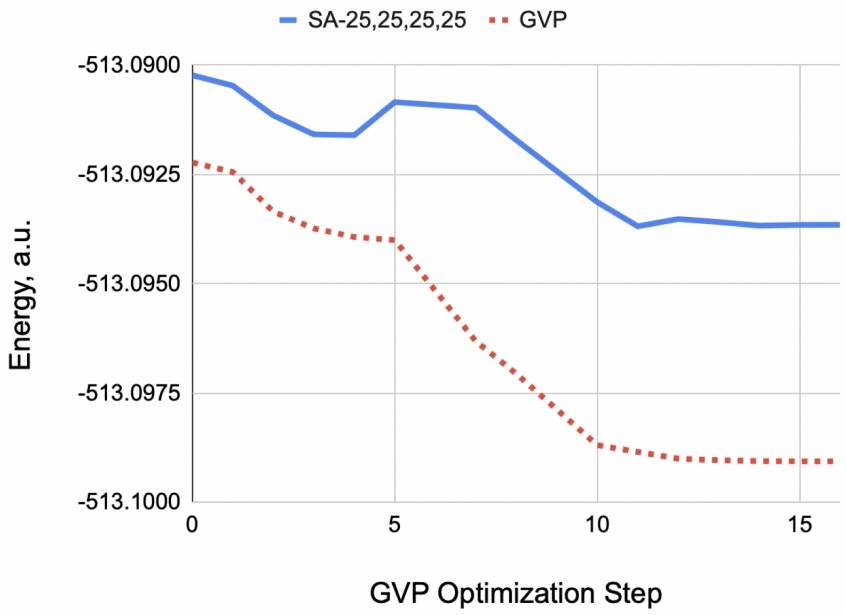


Figure 3.7. Potential energy surface plots of thioacrolein across the GVP Method A geometry optimization steps using GVP (dashed) and equal weighted SA (solid).

state, [121] with SA-CASSCF and SS-CASSCF producing qualitatively different potential energy surfaces. In the case of ABN, the SS result agreed with experiment while the SA result did not. Here, we are again dealing with a state with significant CT character (basically, from a sulfur lone pair towards the ethyl π^*), and so we anticipate that SS-CASSCF, with its ability to fully relax the shapes of the non-active orbitals in response to the CT, is again making the more trustworthy geometric prediction. This perspective is reinforced by the smooth nature of the SS-CASSCF geometry optimization, in contrast to the discontinuities encountered by the biased-weight SA results.

3.5 Conclusion

We have applied both SS- and SA-CASSCF to excited states of thioacrolein to compare and contrast their behavior in a setting where both doubly excited and CT character matter. In an interpolation that produced an avoided crossing, we saw that different computational protocols for SS-CASSCF led to either adiabatic or diabatic potential energy curves, while the presence or absence of discontinuities in SA-CASSCF's curves was determined by the SA weights and was related to both root flipping and discontinuous orbital swapping. When performing unconstrained geometry optimizations, SS-CASSCF and SA-CASSCF agreed about one state's geometry but disagreed qualitatively about another, with SA-CASSCF showing a local minimum at a more planar geometry that was not present in SS-CASSCF. We discussed ways in which these differences mirror previous observations in ABN, where

the more planar SA geometry appeared to be an artifact of its incomplete post-CT orbital relaxation.

There are a number of paths forward for further enhancing the utility of GVP-based SS-CASSCF. Incorporating more properties in the GVP's deviation vector, such as dipoles and overlaps with other states, may give more fine-grained control of whether the method will find diabatic, adiabatic, or spurious energy stationary points, an issue whose importance has recently become more clear [33] and that is further emphasized by the adiabatic and diabatic stationary points found in thioacrolein. One wonders whether the SA-CASSCF step in our Method A could be obviated by the inclusion of overlaps with nearby states in the GVP. Another priority is the evaluation via non-orthogonal CI of interstate properties like transition dipoles and derivative couplings between SS-CASSCF states, as well as the matrix elements needed for re-diagonalization of smooth diabatic surfaces to produce smooth adiabats. Given the advantages that SS-CASSCF has demonstrated both here and in other recent studies, such improvements should offer benefits in many photochemical applications.

3.6 Acknowledgements

This work was supported by the National Science Foundation's CAREER program under Award Number 1848012. Calculations were performed using the Berkeley Research Computing Savio cluster and the Lawrence Berkeley National Lab Lawrencium cluster. R.H. acknowledges that this material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 2146752. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Chapter 4

Conclusion

In this dissertation we have presented a novel approach to excited-state-specific optimization of a complete active space wave function through the minimization of a generalized variational principle. Chapter 2 detailed the formulation of this theory, in which the GVP objective function contains the usual energy gradients along with an additional guiding term which allows for optimization of a broad variety of approximately known properties that help to uniquely pick out the state of interest from among the many possible energy stationary points. The flexibility afforded by the GVP allowed the resolution of states in several challenging cases involving near degeneracies, double excitations, and charge transfer, and located some states that existing state-specific methods have been unable to converge to. In Chapter 3, we explored how this approach affects potential energy surfaces, finding that it can produce smooth diabats or adiabats depending on the computational protocol and that its full state-specific orbital relaxations sometimes produce qualitatively different post-excitation geometry relaxations than those predicted by state averaging.

Unlike SA or existing SS approaches, optimization of the CASSCF ansatz via the GVP in this work utilized a direct one-step approach in which the molecular system's CI coefficients and orbital shapes were optimized simultaneously for each excited state by quasi-Newton minimization. Through this direct approach, we demonstrated that such a parameterization fully overcomes the root flipping problem which continues to be a frustrating reality for two-step methods, and is capable of achieving tighter overall convergence. In addition, by using BFGS and employing an approximate objective function Hessian, this method achieved a computational efficiency competitive with a range of existing SS-CASSCF algorithms. Exploring the advantages of our GVP-CASSCF approach in the context of a geometry relaxations in Chapter 3, we saw the importance of state-specific orbital shapes in the photo-induced C-C bond twist of thioacrolein. Not only did GVP-CASSCF and SA-CASSCF qualitatively disagree on the final geometry, but SA struggled with discontinuities whereas GVP produced a smooth potential energy surface.

Across both studies included in this thesis, we have seen GVP-CASSCF reliably and precisely locate energy stationary points with a robust optimization and provide state-specific orbital descriptions for valence, charge transfer and doubly excited states. This versatility

across multiple classes of excited states is typically not attainable in state averaged approaches. Within the wider context of electronic structure, GVP-CASSCF can serve as a systematically improvable and broadly reliable tool for studying challenging excited states, and can act as a foundation upon which a hierarchy of state-specific post-CASSCF methods can be built.

Appendix A

Supplementary material for the development of GVP-CASSCF

A1 Orbital Energy Derivatives

For the orbital block of the energy derivatives, we define $E^-_{pq} = \left(\hat{a}^\dagger_p \hat{a}_q - \hat{a}^\dagger_q \hat{a}_p\right)$ and $P_{pq,rs}$ as a permutation operator giving us the following expressions for the orbital energy gradient and Hessian.

$$\frac{\partial E}{\partial X_{pq}} = \langle \Psi | \left[E^-_{pq}, \hat{H} \right] | \Psi \rangle \tag{A1}$$

$$\frac{\partial^2 E}{\partial X_{pq} \partial X_{rs}} = \frac{1}{2}(1 + P_{pq,rs}) \langle \Psi | \left[E^-_{pq}, \left[E^-_{rs}, \hat{H} \right] \right] | \Psi \rangle \tag{A2}$$

Core orbitals are indexed using i, j, k, active orbitals with t, u, v, w, and virtual with a, b, c where p, q, r, s are used for general orbitals. For simplicity, we define several Fock-type matrices:

$$F^{core}_{pq} = h_{pq} + \sum_{k}^{N_{core}} \left[2(pq|kk) - (pk|kq) \right] \tag{A3}$$

$$F^{act}_{pq} = \sum_{uv}^{N_{act}} \gamma_{uv} \left[2(pq|uv) - (pu|vq) \right] \tag{A4}$$

$$F^{occ}_{pq} = h_{pq} + \sum_{r}^{N_{occ}} \left[2(pq|rr) - (pr|rq) \right] \tag{A5}$$

where the one and two-electron spin-summed reduced density matrices are defined as

$$\begin{aligned} \gamma_{pq} &= \sum_{IJ} c_I c_J \langle \phi_I | \hat{a}^\dagger_p \hat{a}_q | \phi_J \rangle \\ &= \sum_{IJ} c_I c_J \langle \phi_I | \left(\hat{a}^\dagger_{p_\alpha} \hat{a}_{q_\alpha} + \hat{a}^\dagger_{p_\beta} \hat{a}_{q_\beta} \right) | \phi_J \rangle \end{aligned} \tag{A6}$$

$$\begin{aligned} \Gamma_{pqrs} &= \sum_{IJ} c_I c_J \langle \phi_I | \hat{a}^\dagger_p \hat{a}^\dagger_r \hat{a}_s \hat{a}_q | \phi_J \rangle \\ &= \sum_{IJ} c_I c_J \langle \phi_I | \left(\hat{a}^\dagger_{p_\alpha} \hat{a}^\dagger_{r_\alpha} \hat{a}_{s_\alpha} \hat{a}_{q_\alpha} + \hat{a}^\dagger_{p_\alpha} \hat{a}^\dagger_{r_\beta} \hat{a}_{s_\beta} \hat{a}_{q_\beta} + \hat{a}^\dagger_{p_\alpha} \hat{a}^\dagger_{r_\beta} \hat{a}_{s_\beta} \hat{a}_{q_\alpha} + \hat{a}^\dagger_{p_\beta} \hat{a}^\dagger_{r_\alpha} \hat{a}_{s_\alpha} \hat{a}_{q_\beta} \right) | \phi_J \rangle . \end{aligned} \tag{A7}$$

A1.1 Orbital Energy Gradient

Using the index definitions and Fock-type matrices defined in the previous section, the exact expressions for the core-virtual, active-virtual, and core-active blocks of the orbital energy

gradient in Eq. (A1) evaluated at $X = 0$ are as follows.

$$\frac{\partial E}{\partial X_{ia}} = 4F^{core}_{ai} + 2F^{act}_{ai} \tag{A8}$$

$$\frac{\partial E}{\partial X_{ta}} = 2\sum_{u}^{N_{act}} \gamma_{tu} F^{core}_{au} + 2\sum_{uvw}^{N_{act}} \Gamma_{tuvw}(au|vw) \tag{A9}$$

$$\frac{\partial E}{\partial X_{it}} = 4F^{core}_{ti} + 2F^{act}_{ti} - 2\sum_{u}^{N_{act}} \gamma_{tu} F^{core}_{iu} - 2\sum_{uvw}^{N_{act}} \Gamma_{tuvw}(iu|vw) \tag{A10}$$

A1.2 Approximate Orbital Energy Hessian

A1.2.1 Exact Diagonal

Taking only the diagonal elements of the energy Hessian, Eq. (A2) simplifies to

$$\frac{\partial^2 E}{\partial X^2_{pq}} = \langle \Psi | \left[E^-_{pq}, \left[E^-_{pq}, \hat{H} \right] \right] | \Psi \rangle . \tag{A11}$$

The following are exact expressions for the core-virtual, active-virtual, and core-active blocks of the diagonal orbital energy Hessian in Eq. (A11) evaluated at $X = 0$:

$$\begin{aligned}\frac{\partial^2 E}{\partial X^2_{ia}} &= 4F^{core}_{aa} + 2F^{act}_{aa} - 4F^{core}_{ii} - 2F^{act}_{ii} \\ &\quad - 4(aa|ii) + 12(ai|ai)\end{aligned} \tag{A12}$$

$$\begin{aligned}\frac{\partial^2 E}{\partial X^2_{ta}} &= 2\gamma_{tt} F^{core}_{aa} - 2\sum_{u}^{N_{act}} \gamma_{tu} F^{core}_{tu} - 2\sum_{uvw}^{N_{act}} \Gamma_{tuvw}(tu|vw) \\ &\quad + 2\sum_{uv}^{N_{act}} \left[\Gamma_{tutv}(au|av) + \Gamma_{tvut}(au|av) + \Gamma_{ttvu}(aa|vu) \right]\end{aligned} \tag{A13}$$

$$\begin{aligned}\frac{\partial^2 E}{\partial X^2_{it}} &= 4F^{core}_{tt} + 2F^{act}_{tt} - 4F^{core}_{ii} - 2F^{act}_{ii} \\ &\quad + 2\gamma_{tt} F^{core}_{ii} - 2\sum_{u}^{N_{act}} \gamma_{tu} F^{core}_{tu} - 2\sum_{uvw}^{N_{act}} \Gamma_{tuvw}(tu|vw) \\ &\quad + 2\sum_{uv}^{N_{act}} \left[\Gamma_{tutv}(ui|iv) + \Gamma_{tvut}(ui|iv) + \Gamma_{ttuv}(uv|ii) \right] \\ &\quad + 4\sum_{u}^{N_{act}} \left[3(ui|ui) - (uu|ii) - 3\gamma_{tu}(ui|ti) + \gamma_{tu}(tu|ii) \right] .\end{aligned} \tag{A14}$$

A1.2.2 Fock-based Approximate Diagonal

Adding an additional layer of approximation, we go dropping the off-diagonal terms of the energy Hessian and approximate the Hamiltonian inside the commutators with the one-electron Fock operator, giving us the following approximation to Eq. (A11).

$$\frac{\partial^2 E}{\partial X_{pq}^2} \approx \langle \Psi | \left[E_{pq}^-, \left[E_{pq}^-, \hat{F} \right] \right] | \Psi \rangle \tag{A15}$$

Building the Fock operator from our CASSCF wave function's one-body density matrix and the effective one-electron integrals:

$$\hat{F} = \sum_{pq} \left(h_{pq} + \sum_r \left[2(pq|rr) - (pr|rq) \right] \right) \hat{a}_p^\dagger \hat{a}_q. \tag{A16}$$

With this approximation we arrive at the approximate expressions for core-virtual, active-virtual, and core-active blocks of the diagonal orbital energy Hessian in Eq. (A15) evaluated at $X = 0$:

$$\frac{\partial^2 E}{\partial X_{ia}^2} \approx 2F_{aa}^{occ} - 2F_{ii}^{occ} \tag{A17}$$

$$\frac{\partial^2 E}{\partial X_{ta}^2} \approx 2F_{aa}^{occ}\gamma_{tt} - 2\sum_u^{N_{act}} F_{tu}^{occ}\gamma_{tu} \tag{A18}$$

$$\frac{\partial^2 E}{\partial X_{it}^2} \approx 2F_{ii}^{occ}\gamma_{tt} + 2F_{tt}^{occ} - 2F_{ii}^{occ} - 2\sum_u^{N_{act}} F_{tu}^{occ}\gamma_{tu}. \tag{A19}$$

Appendix B

Supplementary data on the application of GVP-CASSCF

B1 Additional Data for Chapter 2

Table B1. Energies (E_h) of the first excited state $A^1\Sigma^+$ of LiH at various bond lengths.

R (Å)	FCI	$W\Gamma$	GVP
1.2	-7.8421784	-7.8369774	-7.8379204
1.4	-7.8718929	-7.8685677	-7.8689355
1.6	-7.8873115	-7.8843640	-7.8844385
1.8	-7.8950433	-7.8921683	-7.8930879
2.0	-7.8987095	-7.8958730	-7.8968039
2.2	-7.9002698	-7.8973900	-7.8983689
2.4	-7.9007174	-7.8978058	-7.8982932
2.6	-7.9005042	-7.8975273	-7.8979879
2.8	-7.8997797	-7.8966386	-7.8971273
3.0	-7.8985339	-7.8953840	-7.8957249
3.4	-7.8931780	-7.8908310	-7.8907296
3.8	-7.8879230	-7.8847120	-7.8846122
4.2	-7.8809573	-7.8783253	-7.8782487

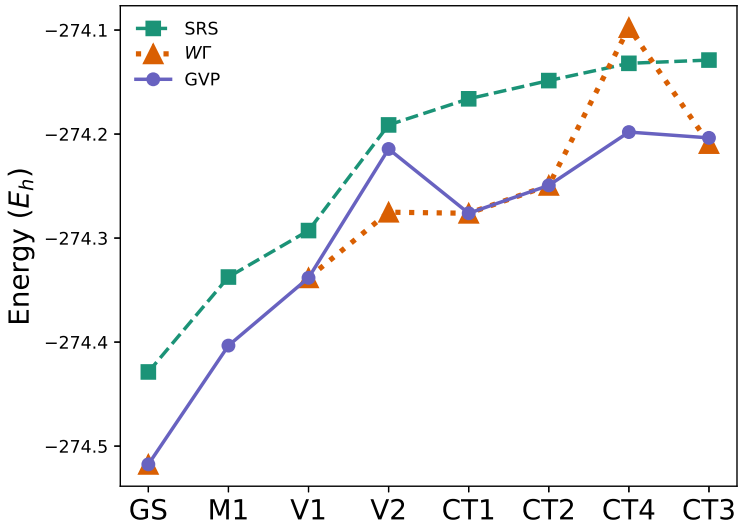


Figure B1. Energy ordering of first eight MgO $^1A_1$ states of the initial CASCI roots and after optimization with the $W\Gamma$ and GVP approaches.

Table B2. Energies (E_h) of the $^1A_1$ states in MgO, listed from top to bottom in ascending order of the CASCI-LDA energies. Labels (GS, M1, etc) are taken from a previous study. [16]

State	Label	CASCI	$W\Gamma$	GVP
1^1A_1	GS	-274.42869956	-274.51755503	-274.51755511
2^1A_1	M1	-274.33744776	–	-274.40336697
3^1A_1	V1	-274.29276479	-274.33820474	-274.33820504
4^1A_1	V2	-274.19120544	-274.27510790	-274.21432010
5^1A_1	CT1	-274.16609490	-274.27614863	-274.27614914
6^1A_1	CT2	-274.14857162	-274.24932368	-274.24932934
7^1A_1	CT4	-274.13197362	-274.09760158	-274.19806809
8^1A_1	CT3	-274.12884711	-274.20910669	-274.20364194

B2 Additional Data for Chapter 3

The following are Z-matrices for the geometry interpolation discussed in Section 3.4.3 and are used as the starting point for unconstrained geometry optimizations in Section 3.4.4.

```
C
C    1    B1
H    1    B2    2    A2
H    1    B3    2    A3    3    D3
H    2    B4    1    A4    3    D4
C    2    B5    1    A5    3    D5
S    6    B6    2    A6    1    D6
H    6    B7    2    A7    1    D7
```

Table B3. Z-matrices for the planar ground state equilibrium geometry of thioacrolein and the slightly twisted geometry found by freezing $D5$ and relaxing the remaining degrees of freedom using SA-1,1,98,0. All bond lengths are reported in Angstroms and angles in degrees.

	Planar	Twisted
B1	1.34839695	1.48305303
B2	1.08172632	1.08069323
A2	121.53657477	119.68009706
B3	1.08179995	1.08085009
A3	120.70521824	119.24918950
D3	179.99484214	159.11396415
B4	1.08446282	1.08188697
A4	119.10580261	119.27932405
D4	-0.00456311	9.49031647
B5	1.47170152	1.41149053
A5	125.65897203	122.50018371
D5	180.00108713	185.00000000
B6	1.62995597	1.78727703
A6	127.24481007	125.07628024
D6	0.00940768	1.34720044
B7	1.08705405	1.08286925
A7	114.50718773	118.41836529
D7	180.00855445	179.92462035

Bibliography

[1] Maylis Orio, Dimitrios A Pantazis, and Frank Neese. "Density functional theory". In: *Photosynth. Res.* 102 (2009), pp. 443–453.

[2] Leszek Lapinski et al. "A Bistable Molecular Switch Driven by Photoinduced Hydrogen-Atom Transfer". In: *ChemPhysChem* 10.13 (2009), pp. 2290–2295.

[3] Tomáš Polívka and Villy Sundström. "Ultrafast dynamics of carotenoid excited states—from solution to natural and artificial systems". In: *Chem. Rev.* 104.4 (2004), pp. 2021–2072. ISSN: 0009-2665.

[4] James F Harrison. "Electronic structure of diatomic molecules composed of a first-row transition metal and main-group element (H-F)". In: *Chem. Rev.* 100.2 (2000), pp. 679–716. ISSN: 0009-2665.

[5] Massimo Olivucci. *Computational photochemistry*. Elsevier, 2005.

[6] Soumen Ghosh et al. "Combining wave function methods with density functional theory for excited states". In: *Chem. Rev.* 118.15 (2018), pp. 7249–7292.

[7] Hans Lischka et al. "Multireference approaches for excited states of molecules". In: *Chem. Rev.* 118.15 (2018), pp. 7293–7361.

[8] WJ Glover. "Communication: Smoothing out excited-state dynamics: Analytical gradients for dynamically weighted complete active space self-consistent field". In: *J. Chem. Phys.* 141.17 (2014).

[9] WJ Glover et al. "Analytical gradients and derivative couplings for dynamically weighted complete active space self-consistent field". In: *J. Chem. Phys.* 151.20 (2019).

[10] Jie Yang et al. "Imaging CF3I conical intersection and photodissociation dynamics with ultrafast electron diffraction". In: *Science* 361.6397 (2018), pp. 64–67.

[11] Patrick M Hare, Carlos E Crespo-Hernández, and Bern Kohler. "Internal conversion to the electronic ground state occurs via two distinct pathways for pyrimidine bases in aqueous solution". In: *Proc. Natl. Acad. Sci. U.S.A.* 104.2 (2007), pp. 435–440.

[12] Péter G Szalay et al. "Benchmark studies on the building blocks of DNA. 2. Effect of biological environment on the electronic excitation spectrum of nucleobases". In: *J. Phys. Chem. A* 116.35 (2012), pp. 8851–8860.

[13] Patrick M Hare, Carlos E Crespo-Hernandez, and Bern Kohler. "Solvent-dependent photophysics of 1-cyclohexyluracil: ultrafast branching in the initial bright state leads nonradiatively to the electronic ground state and a long-lived $1n\pi^*$ state". In: *J. Phys. Chem. B* 110.37 (2006), pp. 18641–18650.

[14] Edward I Solomon et al. "Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins". In: *Chem. Rev.* 104.2 (2004), pp. 419–458.

[15] Carolin König and Johannes Neugebauer. "Quantum chemical description of absorption properties and excited-state processes in photosynthetic systems". In: *ChemPhysChem* 13.2 (2012), pp. 386–425.

[16] L. N. Tran, J. A. R. Shea, and E. Neuscamman. "Tracking excited states in wave function optimization using density matrices and variational principles". In: *J. Chem. Theory Comput.* 15.9 (2019), pp. 4790–4803.

[17] D. A. Kreplin, P. J. Knowles, and H. J. Werner. "Second-order MCSCF optimization revisited. I. Improved algorithms for fast and robust second-order CASSCF convergence". In: *J. Chem. Phys.* 150.19 (2019). ISSN: 0021-9606. DOI: 10.1063/1.5094644. URL: %3CGo%20to%20ISI%3E://WOS:000473298900009.

[18] Michael P Deskevich, David J Nesbitt, and Hans-Joachim Werner. "Dynamically weighted multiconfiguration self-consistent field: Multistate calculations for F^+ H_2O $\rightarrow$ HF^+ OH reaction paths". In: *J. Chem. Phys.* 120.16 (2004), pp. 7281–7289. ISSN: 0021-9606.

[19] Leon Otis, Isabel Craig, and Eric Neuscamman. "A hybrid approach to excited-state-specific variational Monte Carlo and doubly excited states". In: *J. Chem. Phys.* 153.23 (2020), p. 234105. ISSN: 0021-9606.

[20] Robert E Thomas et al. "Stochastic multiconfigurational self-consistent field theory". In: *J. Chem. Theory Comput.* 11.11 (2015), pp. 5316–5325. ISSN: 1549-9618.

[21] J. A. R. Shea, E. Gwin, and E. Neuscamman. "A generalized variational principle with applications to excited state mean field theory". In: *J. Chem. Theory Comput.* 16.3 (2020), pp. 1526–1540. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.9b01105. URL: %3CGo%20to%20ISI%3E://WOS:000519337700015.

[22] Jeppe Olsen et al. "Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces". In: *J. Chem. Phys.* 89 (1988), p. 2185.

[23] Giovanni Li Manni et al. "SplitGAS method for strong correlation and the challenging case of Cr2". In: *J. Chem. Theory Comput.* 9.8 (2013), pp. 3375–3384. ISSN: 1549-9618.

[24] Dominika Zgid and Marcel Nooijen. "The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space". In: *J. Chem. Phys.* 128.14 (2008), p. 144116. ISSN: 0021-9606.

[25] James ET Smith et al. "Cheap and near exact CASSCF with large active spaces". In: *J. Chem. Theory Comput.* 13.11 (2017), pp. 5468–5478. ISSN: 1549-9618.

[26] Giovanni Li Manni, Simon D Smart, and Ali Alavi. "Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte Carlo within a super-CI framework, with application to challenging metal-porphyrins". In: *J. Chem. Theory Comput.* 12.3 (2016), pp. 1245–1258. ISSN: 1549-9618.

[27] Sandeep Sharma et al. "Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory". In: *J. Chem. Theory Comput.* 13.4 (2017), pp. 1595–1604. ISSN: 1549-9618.

[28] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.

[29] T. Helgaker, P. Jørgensen, and J. Olsen. In: *Molecular Electronic Structure Theory*. John Wily and Sons, Ltd: West Sussex, U.K., 2000, pp. 600–610.

[30] John C Tully. "Perspective: Nonadiabatic dynamics theory". In: *J. Chem. Phys.* 137.22 (2012).

[31] Andrei Tokmakoff. *Time-dependent quantum mechanics and spectroscopy*. 2014.

[32] A Boutalib and FX Gadea. "Ab initio adiabatic and diabatic potential-energy curves of the LiH molecule". In: *J. Chem. Phys.* 97.2 (1992), pp. 1144–1156.

[33] Antoine Marie and Hugh GA Burton. "Excited states, symmetry breaking, and unphysical solutions in state-specific CASSCF theory". In: *J. Phys. Chem. A* 127.20 (2023), pp. 4538–4552.

[34] Troy Van Voorhis et al. "The diabatic picture of electron transfer, reaction barriers, and molecular dynamics". In: *Annu. Rev. Phys. Chem.* 61 (2010), pp. 149–170.

[35] Uwe Manthe and H Köppel. "Dynamics on potential energy surfaces with a conical intersection: Adiabatic, intermediate, and diabatic behavior". In: *J. Chem. Phys.* 93.3 (1990), pp. 1658–1669.

[36] Joseph E Subotnik et al. "The requisite electronic structure theory to describe photoexcited nonadiabatic dynamics: Nonadiabatic derivative couplings and diabatic electronic couplings". In: *Acc. Chem. Res.* 48.5 (2015), pp. 1340–1350.

[37] Kate K. Docken and Juergen Hinze. "LiH Potential Curves and Wavefunctions for $X^1\Sigma^+$, $A^1\Sigma^+$, $B^1\Pi$, $^3\Sigma^+$, and $^3\Pi$". In: *J. Chem. Phys.* 57.11 (1972), pp. 4928–4936. DOI: 10.1063/1.1678164. URL: https://aip.scitation.org/doi/abs/10.1063/1.1678164.

[38] David R Yarkony. "Conical intersections: Diabolical and often misunderstood". In: *Acc. Chem. Res.* 31.8 (1998), pp. 511–518.

[39] Walter Kohn, Axel D Becke, and Robert G Parr. "Density functional theory of electronic structure". In: *J. Phys. Chem.* 100.31 (1996), pp. 12974–12980.

[40] Narbe Mardirossian and Martin Head-Gordon. "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals". In: *Mol. Phys.* 115.19 (2017), pp. 2315–2372.

[41] Peter J Knowles and Nicholas C Handy. "A new determinant-based full configuration interaction method". In: *Chem. Phys. Lett.* 111.4-5 (1984), pp. 315–321.

[42] Rodney J Bartlett and Monika Musiał. "Coupled-cluster theory in quantum chemistry". In: *Rev. Mod. Phys.* 79.1 (2007), p. 291.

[43] Clemens Carel Johannes Roothaan. "New developments in molecular orbital theory". In: *Rev. Mod. Phys.* 23.2 (1951), p. 69.

[44] George G Hall. "The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials". In: *Proc. R. Soc. A: Math. Phys. Eng. Sci.* 205.1083 (1951), pp. 541–552.

[45] Rodney J Bartlett and John F Stanton. "Applications of Post-Hartree—Fock Methods: A Tutorial". In: *Rev. Comput. Chem.* (1994), pp. 65–169.

[46] R. D. Johnson. *NIST Computational Chemistry Comparison and Benchmark Database.* Online Database. 2019. URL: `http://cccbdb.nist.gov/`.

[47] Rodney J Bartlett. "Many-body perturbation theory and coupled cluster theory for electron correlation in molecules". In: *Annu. Rev. Phys. Chem.* 32.1 (1981), pp. 359–401.

[48] Klaus Ruedenberg et al. "Are atoms intrinsic to molecular electronic wavefunctions? I. The FORS model". In: *Chem. Phys.* 71.1 (1982), pp. 41–49. ISSN: 0301-0104.

[49] Hans-Joachim Werner and Peter J Knowles. "A second order multiconfiguration SCF procedure with optimum convergence". In: *J. Chem. Phys.* 82.11 (1985), pp. 5053–5063. ISSN: 0021-9606.

[50] Peter J Knowles and Hans-Joachim Werner. "An efficient second-order MC SCF method for long configuration expansions". In: *Chem. Phys. Lett.* 115.3 (1985), pp. 259–267. ISSN: 0009-2614.

[51] Björn O Roos. "The complete active space self-consistent field method and its applications in electronic structure calculations". In: *Adv. Chem. Phys.* 69 (1987), pp. 399–445. ISSN: 0065-2385.

[52] Björn O Roos, Peter R Taylor, and Per EM Sigbahn. "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach". In: *Chem. Phys.* 48.2 (1980), pp. 157–173. ISSN: 0301-0104.

[53] K. Ruedenberg, L. M. Cheung, and S. T. Elbert. "MCSCF optimization through combined use of natural orbitals and the brillouin–levy–berthier theorem". In: *Int. J. Quantum Chem.* 16.5 (1979), pp. 1069–1101. ISSN: 0020-7608. DOI: https://doi.org/10.1002/qua.560160511. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.560160511.

[54] Konstantinos D Vogiatzis et al. "Pushing configuration-interaction to the limit: Towards massively parallel MCSCF calculations". In: *J. Chem. Phys.* 147.18 (2017).

[55] Debashree Ghosh et al. "Orbital optimization in the density matrix renormalization group, with applications to polyenes and β-carotene". In: *J. Chem. Phys.* 128.14 (2008), p. 144117. ISSN: 0021-9606.

[56] Christopher J Stein and Markus Reiher. "Automated selection of active orbital spaces". In: *J. Chem. Theory Comput.* 12.4 (2016), pp. 1760–1771.

[57] Kerstin Andersson et al. "Second-order perturbation theory with a CASSCF reference function". In: *J. Phys. Chem.* 94.14 (1990), pp. 5483–5488.

[58] Kerstin Andersson, Per-Åke Malmqvist, and Björn O Roos. "Second-order perturbation theory with a complete active space self-consistent field reference function". In: *J. Chem. Phys.* 96.2 (1992), pp. 1218–1226.

[59] Danny L Yeager et al. "Newton-Raphson approaches and generalizations in multiconfigurational self-consistent field calculations". In: *J. Phys. Chem.* 86.12 (1982), pp. 2140–2153.

[60] Hugh GA Burton. "Energy Landscape of State-Specific Electronic Structure Theory". In: *J. Chem. Theory Comput.* 18.3 (2022), pp. 1512–1526. ISSN: 1549-9618.

[61] Mathieu Lewin. "On the computation of excited states with MCSCF methods". In: *J. Math. Chem.* 44.4 (2008), pp. 967–980.

[62] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge university press, 2018.

[63] Garnet Kin-Lic Chan and Sandeep Sharma. "The density matrix renormalization group in quantum chemistry". In: *Annu. Rev. Phys. Chem.* 62 (2011), pp. 465–481.

[64] Randall N Diffenderfer and David R Yarkony. "Use of the state-averaged MCSCF procedure: application to radiative transitions in magnesium oxide". In: *J. Phys. Chem.* 86.26 (1982), pp. 5098–5105.

[65] J. Gavnholt et al. "Delta self-consistent field method to obtain potential energy surfaces of excited molecules on surfaces". In: *Phys. Rev. B* 78.7 (2008). ISSN: 2469-9950. DOI: 10.1103/PhysRevB.78.075441. URL: %3CGo%20to%20ISI%3E://WOS:000259406700122.

[66] Giuseppe MJ Barca, Andrew TB Gilbert, and Peter MW Gill. "Simple models for difficult electronic excitations". In: *J. Chem. Theory Comput.* 14.3 (2018), pp. 1501–1509. ISSN: 1549-9618.

[67] Andrew TB Gilbert, Nicholas A Besley, and Peter MW Gill. "Self-consistent field calculations of excited states using the maximum overlap method (MOM)". In: *J. Phys. Chem. A* 112.50 (2008), pp. 13164–13171. ISSN: 1089-5639.

[68] Jan-M Mewes et al. "On the molecular mechanism of non-radiative decay of nitrobenzene and the unforeseen challenges this simple molecule holds for electronic structure theory". In: *Phys. Chem. Chem. Phys.* 16.24 (2014), pp. 12393–12406.

[69] Gianluca Levi, Aleksei V Ivanov, and Hannes Jónsson. "Variational density functional calculations of excited states via direct optimization". In: *J. Chem. Theory Comput.* 16.11 (2020), pp. 6968–6982.

[70] Nicholas A Besley, Andrew TB Gilbert, and Peter MW Gill. "Self-consistent-field calculations of core excited states". In: *J. Chem. Phys.* 130.12 (2009).

[71] R. P. Messmer. "On a variational method for determining excited state wave functions". In: *Theor. Chim. Acta* 14 (1969), pp. 319–328.

[72] J. H. Choi, C. F. Lebeda, and R. P. Messmer. "Variational Principle for excited states: Exact formulation and other extensions". In: *Chem. Phys. Lett.* 5 (1970), pp. 503–506.

[73] D. Hait and M. Head-Gordon. "Excited State Orbital Optimization via Minimizing the Square of the Gradient: General Approach and Application to Singly and Doubly Excited States via Density Functional Theory". In: *J. Chem. Theory Comput.* 16.3 (2020), pp. 1699–1710. DOI: 10.1021/acs.jctc.9b01127.

[74] Troy Van Voorhis and Martin Head-Gordon. "A geometric approach to direct minimization". In: *Mol. Phys.* 100.11 (2002), pp. 1713–1721. ISSN: 0026-8976.

[75] Francesco Aquilante et al. "Accurate ab initio density fitting for multiconfigurational self-consistent field methods". In: *J. Chem. Phys.* 129.2 (2008), p. 024113. DOI: 10.1063/1.2953696. URL: https://aip.scitation.org/doi/abs/10.1063/1.2953696.

[76] Edward G. Hohenstein et al. "An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units". In: *J. Chem. Phys.* 142.22 (2015), p. 224103. DOI: 10.1063/1.4921956. URL: https://aip.scitation.org/doi/abs/10.1063/1.4921956.

[77] Björn O. Roos. "The complete active space SCF method in a fock-matrix-based super-CI formulation". In: *Int. J. Quantum Chem.* 18.S14 (1980), pp. 175–189. ISSN: 0020-7608. DOI: https://doi.org/10.1002/qua.560180822. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.560180822.

[78] Per E. M. Siegbahn et al. "The complete active space SCF (CASSCF) method in a Newton–Raphson formulation with application to the HNO molecule". In: *J. Chem. Phys.* 74.4 (1981), pp. 2384–2396. DOI: 10.1063/1.441359. URL: https://aip.scitation.org/doi/abs/10.1063/1.441359.

[79] Danny L. Yeager and Poul Jørgensen. "Convergency studies of second and approximate second order multiconfigurational Hartree–Fock procedures". In: *J. Chem. Phys.* 71.2 (1979), pp. 755–760. DOI: 10.1063/1.438363. URL: https://aip.scitation.org/doi/abs/10.1063/1.438363.

[80] Yousef Saad. *Numerical methods for large eigenvalue problems: revised edition.* SIAM, 2011.

[81] I Shavitt et al. "The iterative calculation of several of the lowest or highest eigenvalues and corresponding eigenvectors of very large symmetric matrices". In: *J. Comput. Phys.* 11.1 (1973), pp. 90–108. ISSN: 0021-9991.

[82] Neculai Andrei. "Scaled conjugate gradient algorithms for unconstrained optimization". In: *Comput. Optim. Appl.* 38.3 (2007), pp. 401–416. ISSN: 1573-2894.

[83] Per EM Siegbahn et al. "A comparison of the super-CI and the Newton-Raphson scheme in the complete active space SCF method". In: *Phys. Scripta* 21.3-4 (1980), p. 323. ISSN: 1402-4896.

[84] D. A. Kreplin, P. J. Knowles, and H. J. Werner. "MCSCF optimization revisited. II. Combined first- and second-order orbital optimization for large molecules". In: *J. Chem. Phys.* 152.7 (2020). ISSN: 0021-9606. DOI: 10.1063/1.5142241. URL: %3CGo%20to%20ISI%3E://WOS:000519820500001.

[85] Q. M. Sun, J. Yang, and G. K. L. Chan. "A general second order complete active space self-consistent-field solver for large-scale systems". In: *Chem. Phys. Lett.* 683 (2017), pp. 291–299. ISSN: 0009-2614. DOI: 10.1016/j.cplett.2017.03.004. URL: %3CGo%20to%20ISI%3E://WOS:000405802200047.

[86] Thomas Bondo Pedersen. "Introduction to Response Theory". In: *Handbook of Computational Chemistry.* Ed. by Jerzy Leszczynski. Dordrecht: Springer Netherlands, 2012, pp. 135–156. ISBN: 978-94-007-0711-5. DOI: 10.1007/978-94-007-0711-5_5. URL: https://doi.org/10.1007/978-94-007-0711-5_5.

[87] Anna I Krylov. "Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space". In: *Annu. Rev. Phys. Chem.* 59 (2008), pp. 433–462.

[88] Piotr Piecuch, Jared A Hansen, and Adeayo O Ajala. "Benchmarking the completely renormalised equation-of-motion coupled-cluster approaches for vertical excitation energies". In: *Mol. Phys.* 113.19-20 (2015), pp. 3085–3127.

[89] John Sous, Prateek Goel, and Marcel Nooijen. "Similarity transformed equation of motion coupled cluster theory revisited: a benchmark study of valence excited states". In: *Mol. Phys.* 112.5-6 (2014), pp. 616–638.

[90] Varun Rishi et al. "Excited states from modified coupled cluster methods: Are they any better than EOM CCSD?" In: *J. Chem. Phys.* 146.14 (2017).

[91] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2013.

[92] Jesse J Lutz et al. "Reference dependence of the two-determinant coupled-cluster method for triplet and open-shell singlet states of biradical molecules". In: *J. Chem. Phys.* 148.16 (2018).

[93] Péter G Szalay et al. "Benchmark studies on the building blocks of DNA. 3. Watson–Crick and stacked base pairs". In: *J. Phys. Chem. A* 117.15 (2013), pp. 3149–3157.

[94] Mark E Casida and Miquel Huix-Rotllant. "Progress in time-dependent density-functional theory". In: *Annu. Rev. Phys. Chem.* 63 (2012), pp. 287–323.

[95] Andreas Dreuw and Martin Head-Gordon. "Single-reference ab initio methods for the calculation of excited states of large molecules". In: *Chem. Rev.* 105.11 (2005), pp. 4009–4037.

[96] Neepa T Maitra et al. "Double excitations within time-dependent density functional theory linear response". In: *J. Chem. Phys.* 120.13 (2004), pp. 5932–5937.

[97] Benjamin G Levine et al. "Conical intersections and double excitations in time-dependent density functional theory". In: *Mol. Phys.* 104.5-7 (2006), pp. 1039–1051.

[98] Andreas Dreuw, Jennifer L Weisman, and Martin Head-Gordon. "Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange". In: *J. Chem. Phys.* 119.6 (2003), pp. 2943–2946.

[99] Andreas Dreuw and Martin Head-Gordon. "Failure of time-dependent density functional theory for long-range charge-transfer excited states: the zincbacteriochlorin-bacteriochlorin and bacteriochlorophyll- spheroidene complexes". In: *J. Am. Chem. Soc.* 126.12 (2004), pp. 4007–4016.

[100] Peter Elliott et al. "Perspectives on double-excitations in TDDFT". In: *Chem. Phys.* 391.1 (2011), pp. 110–119.

[101] David J Tozer and Nicholas C Handy. "On the determination of excitation energies using density functional theory". In: *Phys. Chem. Chem. Phys.* 2.10 (2000), pp. 2117–2121.

[102] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. "Challenges for density functional theory". In: *Chem. Rev.* 112.1 (2012), pp. 289–320.

[103] Gary S Kedziora et al. "Bond breaking in stretched molecules: multi-reference methods versus density functional theory". In: *Theor. Chem. Acc.* 135.3 (2016), p. 79.

[104] Roi Baer, Ester Livshits, and Ulrike Salzner. "Tuned range-separated hybrids in density functional theory". In: *Annu. Rev. Phys. Chem.* 61 (2010), pp. 85–109.

[105] J. A. R. Shea and E. Neuscamman. "Communication: A mean field platform for excited state quantum chemistry". In: *J. Chem. Phys.* 149.8 (2018). ISSN: 0021-9606. DOI: 10.1063/1.5045056. URL: %3CGo%20to%20ISI%3E://WOS:000444035800002.

[106] R. Clune, J. A. R. Shea, and E. Neuscamman. "N-5-scaling excited-state-specific perturbation theory". In: *J. Chem. Theory Comput.* 16.10 (2020), pp. 6132–6141. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.0c00308. URL: %3CGo%20to%20ISI%3E://WOS:000580954000014.

[107] Harrison Tuckman and Eric Neuscamman. "An Excited-State-Specific Projected Coupled-Cluster Theory". In: *arXiv preprint arXiv:2302.06731* (2023).

[108] T. S. Hardikar and E. Neuscamman. "A self-consistent field formulation of excited state mean field theory". In: *J. Chem. Phys.* 153.16 (2020). ISSN: 0021-9606. DOI: 10.1063/5.0019557. URL: %3CGo%20to%20ISI%3E://WOS:000587324700004.

[109] Luning Zhao and Eric Neuscamman. "Excited state mean-field theory without automatic differentiation". In: *J. Chem. Phys.* 152.20 (2020), p. 204112. ISSN: 0021-9606.

[110] Rachel Clune et al. "Studying excited-state-specific perturbation theory on the Thiel set". In: *J. Chem. Phys.* 158.22 (2023).

[111] Hans-Joachim Werner and Wilfried Meyer. "A quadratically convergent MCSCF method for the simultaneous optimization of several states". In: *J. Chem. Phys.* 74.10 (1981), pp. 5794–5801. ISSN: 0021-9606.

[112] Frank Neese et al. "Advanced aspects of ab initio theoretical optical spectroscopy of transition metal complexes: Multiplets, spin-orbit coupling and resonance Raman intensities". In: *Coord. Chem. Rev.* 251.3-4 (2007), pp. 288–327.

[113] Ken Tokunaga, Tohru Sato, and Kazuyoshi Tanaka. "Calculation of vibronic coupling constant and vibronic coupling density analysis". In: *J. Mol. Struct.* 838.1-3 (2007), pp. 116–123.

[114] Ignacio Fdez. Galván et al. "Analytical state-average complete-active-space self-consistent field nonadiabatic coupling vectors: Implementation with density-fitted two-electron integrals and application to conical intersections". In: *J. Chem. Theory Comput.* 12.8 (2016), pp. 3636–3653. ISSN: 1549-9618.

[115] Samer Gozem et al. "Shape of multireference, equation-of-motion coupled-cluster, and density functional theory potential energy surfaces at a conical intersection". In: *J. Chem. Theory Comput.* 10.8 (2014), pp. 3074–3084. ISSN: 1549-9618.

[116] Per-Åke Malmqvist and Björn O Roos. "The CASSCF state interaction method". In: *Chem. Phys. Lett.* 155.2 (1989), pp. 189–194. ISSN: 0009-2614.

[117] Luis Serrano-Andrés, Manuela Merchán, and Roland Lindh. "Computation of conical intersections by using perturbation techniques". In: *J. Chem. Phys.* 122.10 (2005), p. 104107. ISSN: 0021-9606.

[118] Hans Lischka, M Dallos, and R Shepard. "Analytic MRCI gradient for excited states: formalism and application to the $n - \pi^*$ valence and $n-$(3s, 3p) Rydberg states of formaldehyde". In: *Mol. Phys.* 100.11 (2002), pp. 1647–1658. ISSN: 0026-8976.

[119] Jonna Stålring, Anders Bernhardsson, and Roland Lindh. "Analytical gradients of a state average MCSCF state and a state average diagnostic". In: *Mol. Phys.* 99.2 (2001), pp. 103–114. ISSN: 0026-8976.

[120] Mireia Segado, Isabel Gómez, and Mar Reguero. "Intramolecular charge transfer in aminobenzonitriles and tetrafluoro counterparts: fluorescence explained by competition between low-lying excited states and radiationless deactivation. Part I: A mechanistic overview of the parent system ABN". In: *Phys. Chem. Chem. Phys.* 18.9 (2016), pp. 6861–6874.

[121] L. N. Tran and E. Neuscamman. "Improving Excited-State Potential Energy Surfaces via Optimal Orbital Shapes". In: *J. Phys. Chem. A* 124.40 (2020), pp. 8273–8279. ISSN: 1089-5639. DOI: 10.1021/acs.jpca.0c07593. URL: %3CGo%20to%20ISI%3E://WOS:000580563800017.

[122] Peter Pulay. "A perspective on the CASPT2 method". In: *Int. J. Quantum Chem.* 111.13 (2011), pp. 3273–3279.

[123] Shu-Hua Xia et al. "Excited-state ring-opening mechanism of cyclic ketones: a MS-CASPT2//CASSCF study". In: *J. Phys. Chem. A* 119.15 (2015), pp. 3569–3576.

[124] Toshifumi Mori and Shigeki Kato. "Dynamic electron correlation effect on conical intersections in photochemical ring-opening reaction of cyclohexadiene: MS-CASPT2 study". In: *Chem. Phys. Lett.* 476.1-3 (2009), pp. 97–100.

[125] Lihong Liu, Jian Liu, and Todd J Martinez. "Dynamical correlation effects on photoisomerization: Ab initio multiple spawning dynamics with MS-CASPT2 for a model trans-protonated Schiff base". In: *J. Phys. Chem. B* 120.8 (2016), pp. 1940–1949.

[126] Jae Woo Park and Toru Shiozaki. "On-the-fly CASPT2 surface-hopping dynamics". In: *J. Chem. Theory Comput.* 13.8 (2017), pp. 3676–3683.

[127] Iakov Polyak et al. "Ultrafast photoinduced dynamics of 1, 3-cyclohexadiene using XMS-CASPT2 surface hopping". In: *J. Chem. Theory Comput.* 15.7 (2019), pp. 3929–3940.

[128] Sebastian Mai et al. "The influence of the electronic structure method on intersystem crossing dynamics. The case of thioformaldehyde". In: *J. Chem. Theory Comput.* 15.6 (2019), pp. 3470–3480.

[129] Dominikus Brian et al. "Three-state harmonic models for photoinduced charge transfer". In: *J. Chem. Phys.* 154.17 (2021), p. 174105. ISSN: 0021-9606.

[130] Harry A Frank et al. "Effect of the solvent environment on the spectroscopic properties and dynamics of the lowest excited states of carotenoids". In: *J. Phys. Chem. B* 104.18 (2000), pp. 4569–4577. ISSN: 1520-6106.

[131] HM Dhammika Bandara and Shawn C Burdette. "Photoisomerization in different classes of azobenzene". In: *Chem. Soc. Rev.* 41.5 (2012), pp. 1809–1825.

[132] Dario Polli et al. "Conical intersection dynamics of the primary photoisomerization event in vision". In: *Nature* 467.7314 (2010), pp. 440–443. ISSN: 1476-4687.

[133] George Zimmerman, Lue-Yung Chow, and Un-Jin Paik. "The photochemical isomerization of azobenzene1". In: *J. Am. Chem. Soc.* 80.14 (1958), pp. 3528–3531. ISSN: 0002-7863.

[134] Emily E Claveau and Evangelos Miliordos. "Electronic structure of the dicationic first row transition metal oxides". In: *Phys. Chem. Chem. Phys.* (2021).

[135] Evangelos Miliordos and Aristides Mavridis. "Electronic structure and bonding of the early 3d-transition metal diatomic oxides and their ions: ScO, TiO, CrO, and MnO". In: *J. Phys. Chem. A* 114.33 (2010), pp. 8536–8572. ISSN: 1089-5639.

[136] Thomas Bouabça et al. "A study of the fixed-node error in quantum Monte Carlo calculations of electronic transitions: The case of the singlet $n \rightarrow \pi^*$ (CO) transition of the acrolein". In: *J. Chem. Phys.* 130.11 (2009), p. 114107. ISSN: 0021-9606.

[137] Alexander A Granovsky. "Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory". In: *J. Chem. Phys.* 134.21 (2011), p. 214113. ISSN: 0021-9606.

[138] M Chandler Bennett. "High-accuracy electronic structure calculations with QMCPACK". In: *Nat. Rev. Phys.* (2021), pp. 1–1. ISSN: 2522-5820.

[139] Shivesh Pathak et al. "Excited states in variational Monte Carlo using a penalty method". In: *J. Chem. Phys.* 154.3 (2021), p. 034101. ISSN: 0021-9606.

[140] Hong-Zhou Ye et al. "σ-SCF: A direct energy-targeting method to mean-field excited states". In: *J. Chem. Phys.* 147.21 (2017), p. 214104. ISSN: 0021-9606.

[141] Hong-Zhou Ye and Troy Van Voorhis. "Half-projected σ self-consistent field for electronic excited states". In: *J. Chem. Theory Comput.* 15.5 (2019), pp. 2954–2965. ISSN: 1549-9618.

[142] Kevin Carter-Fenk and John M Herbert. "State-targeted energy projection: A simple and robust approach to orbital relaxation of non-aufbau self-consistent field solutions". In: *J. Chem. Theory Comput.* 16.8 (2020), pp. 5067–5082. ISSN: 1549-9618.

[143] Diptarka Hait et al. "Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations". In: *J. Chem. Phys.* 153.13 (2020), p. 134108. ISSN: 0021-9606.

[144] Diptarka Hait and Martin Head-Gordon. "Highly accurate prediction of core spectra of molecules at density functional theory cost: Attaining sub-electronvolt error from a restricted open-shell Kohn–Sham approach". In: *J. Phys. Chem. Lett.* 11.3 (2020), pp. 775–786. ISSN: 1948-7185.

[145] Scott M Garner and Eric Neuscamman. "A variational Monte Carlo approach for core excitations". In: *J. Chem. Phys.* 153.14 (2020), p. 144108. ISSN: 0021-9606.

[146] Scott M Garner and Eric Neuscamman. "Core excitations with excited state mean field and perturbation theory". In: *J. Chem. Phys.* 153.15 (2020), p. 154102. ISSN: 0021-9606.

[147] Fábris Kossoski et al. "Excited States from State-Specific Orbital-Optimized Pair Coupled Cluster". In: *J. Chem. Theory Comput.* 17.8 (2021), pp. 4756–4768. ISSN: 1549-9618.

[148] Antoine Marie, Fábris Kossoski, and Pierre-Francois Loos. "Variational coupled cluster for ground and excited states". In: *J. Chem. Phys.* 155.10 (2021), p. 104105. ISSN: 0021-9606.

[149] Byron H. Lengsfield III. "General second order MCSCF theory: A density matrix directed algorithm". In: *J. Chem. Phys.* 73.1 (1980), pp. 382–390. DOI: 10.1063/1.439885. URL: https://aip.scitation.org/doi/abs/10.1063/1.439885.

[150] Poul Jørgensen, Peter Swanstrøm, and Danny L. Yeager. "Guaranteed convergence in ground state multiconfigurational self-consistent field calculations". In: *J. Chem. Phys.* 78.1 (1983), pp. 347–356. DOI: 10.1063/1.444508. URL: https://aip.scitation.org/doi/abs/10.1063/1.444508.

[151] Per Åke Malmqvist. "Calculation of transition density matrices by nonunitary orbital transformations". In: *Int. J. Quantum Chem.* 30.4 (1986), pp. 479–494. ISSN: 0020-7608. DOI: https://doi.org/10.1002/qua.560300404. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.560300404.

[152] Alex J. W. Thom and Martin Head-Gordon. "Hartree–Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction". In: *J. Chem. Phys.* 131.12 (2009), p. 124113. DOI: 10.1063/1.3236841. URL: https://aip.scitation.org/doi/abs/10.1063/1.3236841.

[153] Eric J. Sundstrom and Martin Head-Gordon. "Non-orthogonal configuration interaction for the calculation of multielectron excited states". In: *J. Chem. Phys.* 140.11 (2014), p. 114103. DOI: 10.1063/1.4868120. URL: https://aip.scitation.org/doi/abs/10.1063/1.4868120.

[154] Charles George Broyden. "The convergence of a class of double-rank minimization algorithms 1. general considerations". In: *IMA J. Appl. Math.* 6.1 (1970), pp. 76–90.

[155] Roger Fletcher. "A new approach to variable metric algorithms". In: *J. Comput.* 13.3 (1970), pp. 317–322.

[156] Donald Goldfarb. "A family of variable-metric methods derived by variational means". In: *Math. Comput.* 24.109 (1970), pp. 23–26.

[157] David F Shanno. "Conditioning of quasi-Newton methods for function minimization". In: *Math. Comput.* 24.111 (1970), pp. 647–656.

[158] Beatrice Van Der Goetz. "Improving Wavefunction Efficiency by Tessellating Correlation Factors and Coupled State-Specific Optimization". PhD thesis. University of California, Berkeley, 2021.

[159] Per Åke Malmqvist, Alistair Rendell, and Björn O Roos. "The restricted active space self-consistent-field method, implemented with a split graph unitary group approach". In: *J. Phys. Chem.* 94.14 (1990), pp. 5477–5482. ISSN: 0022-3654.

[160] Galina Chaban, Michael W Schmidt, and Mark S Gordon. "Approximate second order method for orbital optimization of SCF and MCSCF wavefunctions". In: *Theor. Chem. Acc.* 97.1 (1997), pp. 88–95. ISSN: 1432-2234.

[161] Qiming Sun et al. "PySCF: the Python-based simulations of chemistry framework". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 8.1 (2018), e1340. ISSN: 1759-0876.

[162] Thom H Dunning Jr. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". In: *J. Chem. Phys.* 90.2 (1989), pp. 1007–1023. ISSN: 0021-9606.

[163] Warren J Hehre, Robert Ditchfield, and John A Pople. "Self—consistent molecular orbital methods. XII. Further extensions of Gaussian—type basis sets for use in molecular orbital studies of organic molecules". In: *J. Chem. Phys.* 56.5 (1972), pp. 2257–2261. ISSN: 0021-9606.

[164] Abdul-Rahman Allouche. "Gabedit—A graphical user interface for computational chemistry softwares". In: *J. Comput. Chem.* 32.1 (2011), pp. 174–182. ISSN: 0192-8651. DOI: https://doi.org/10.1002/jcc.21600. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21600.

[165] Gijs Schaftenaar, Elias Vlieg, and Gerrit Vriend. "Molden 2.0: quantum chemistry meets proteins". In: *J. Comput.-Aided Mol. Des.* 31.9 (2017), pp. 789–800. DOI: 10.1007/s10822-017-0042-5.

[166] Ewa Pastorczak, Nikitas I Gidopoulos, and Katarzyna Pernal. "Calculation of electronic excited states of molecules using the Helmholtz free-energy minimum principle". In: *Phys. Rev. A* 87.6 (2013), p. 062501.

[167] Hans Jørgen Aa. Jensen, Poul Jørgensen, and Hans Ågren. "Efficient optimization of large scale MCSCF wave functions with a restricted step algorithm". In: *J. Chem. Phys.* 87.1 (1987), pp. 451–466. DOI: 10.1063/1.453590. URL: https://aip.scitation.org/doi/abs/10.1063/1.453590.

[168] A Maatouk et al. "Electronic states of MgO: Spectroscopy, predissociation, and cold atomic Mg and O production". In: *J. Chem. Phys.* 133.14 (2010), p. 144302. ISSN: 0021-9606.

[169] Jeong Hyun Kim et al. "Vibrationally resolved photoelectron spectroscopy of MgO^- and ZnO^- and the low-lying electronic states of MgO, MgO^-, and ZnO". In: *J. Phys. Chem. A* 105.23 (2001), pp. 5709–5718. ISSN: 1089-5639.

[170] Helmar Thümmel, Rainer Klotz, and Sigrid D Peyerimhoff. "The electronic structure of the MgO molecule in ground and excited states". In: *Chem. Phys.* 129.3 (1989), pp. 417–430. ISSN: 0301-0104.

[171] Randall N Diffenderfer and David R Yarkony. "Use of the state-averaged MCSCF procedure: application to radiative transitions in magnesium oxide". In: *J. Phys. Chem.* 86.26 (1982), pp. 5098–5105. ISSN: 0022-3654.

[172] Kerstin Andersson and Björn O Roos. "Excitation energies in the nickel atom studied with the complete active space SCF method and second-order perturbation theory". In: *Chem. Phys. Lett.* 191.6 (1992), pp. 507–514. ISSN: 0009-2614.

[173] Per Åke Malmqvist et al. "The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems". In: *J. Chem. Phys.* 128.20 (2008), p. 204109. ISSN: 0021-9606.

[174] P. W. Atkins and R. S. Friedman. *Molecular Quantum Mechanics, Third Edition*. New York: Oxford University Press, 1997, p. 362.

[175] Stephan P Velsko and Graham R Fleming. "Photochemical isomerization in solution. Photophysics of diphenyl butadiene". In: *J. Chem. Phys.* 76.7 (1982), pp. 3553–3562.

[176] Benjamin Lasorne, Graham A Worth, and Michael A Robb. "Excited-state dynamics". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 1.3 (2011), pp. 460–475.

[177] David R Yarkony. "Diabolical conical intersections". In: *Rev. Mod. Phys.* 68.4 (1996), p. 985.

[178] Zhongxing Xu et al. "Ab initio study of ground-state CS photodissociation via highly excited electronic states". In: *Astrophys. J.* 882.2 (2019), p. 86.

[179] Run-Ze Mao et al. "Photoinduced C–S Bond Cleavage of Thioglycosides and Glycosylation". In: *Org. Lett.* 17.22 (2015), pp. 5606–5609.

[180] Youssef Abderrazak, Aditya Bhattacharyya, and Oliver Reiser. "Visible-Light-Induced Homolysis of Earth-Abundant Metal-Substrate Complexes: A Complementary Activation Strategy in Photoredox Catalysis". In: *Angew. Chem. Int. Ed.* 60.39 (2021), pp. 21100–21115.

[181] Saumik Sen et al. "Insight into the isomerization mechanism of retinal proteins from hybrid quantum mechanics/molecular mechanics simulations". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 12.1 (2022), e1562.

[182] Marina Blanco-Lomas, Pedro J Campos, and Diego Sampedro. "Synthesis and Photoisomerization of Rhodopsin-Based Molecular Switches". In: *Eur. J. Org. Chem.* 2012.32 (2012), pp. 6328–6334.

[183] Miguel AL Marques and Eberhard KU Gross. "Time-dependent density functional theory". In: *Annu. Rev. Phys. Chem.* 55 (2004), pp. 427–455.

[184] Max Winslow, Warren B Cross, and David Robinson. "Comparison of Spin-Flip TDDFT-Based Conical Intersection Approaches with XMS-CASPT2". In: *J. Chem. Theory Comput.* 16.5 (2020), pp. 3253–3263.

[185] Paolo Celani et al. "Conical intersection mechanism for photochemical ring opening in benzospiropyran compounds". In: *J. Am. Chem. Soc.* 119.44 (1997), pp. 10815–10820.

[186] Fengyi Liu et al. "Multireference ab initio density matrix renormalization group (DMRG)-CASSCF and DMRG-CASPT2 study on the photochromic ring opening of spiropyran". In: *J. Chem. Theory Comput.* 9.10 (2013), pp. 4462–4469.

[187] Todd J Martinez. "Insights for light-driven molecular devices from ab initio multiple spawning excited-state dynamics of organic and biological chromophores". In: *Acc. Chem. Res.* 39.2 (2006), pp. 119–126.

[188] Gareth W Richings et al. "Quantum dynamics simulations using Gaussian wavepackets: the vMCG method". In: *Int. Rev. Phys. Chem.* 34.2 (2015), pp. 269–308.

[189] Piero Altoé et al. "Light driven molecular switches: exploring and tuning their photophysical and photochemical properties". In: *Theor. Chem. Acc.* 117.5-6 (2007), pp. 1041–1059.

[190] Peter J Knowles and Hans-Joachim Werner. "Internally contracted multiconfiguration-reference configuration interaction calculations for excited states". In: *Theor. Chim. Acta* 84 (1992), pp. 95–103.

[191] Weifeng Hu and Garnet Kin-Lic Chan. "Excited-state geometry optimization with the density matrix renormalization group, as applied to polyenes". In: *J. Chem. Theory Comput.* 11.7 (2015), pp. 3000–3009.

[192] Petr Slavíček and Todd J Martínez. "Ab initio floating occupation molecular orbital-complete active space configuration interaction: An efficient approximation to CASSCF". In: *J. Chem. Phys.* 132.23 (2010).

[193] Lan Nguyen Tran and Eric Neuscamman. "Exploring Ligand-to-Metal Charge-Transfer States in the Photo-Ferrioxalate System Using Excited-State Specific Optimization". In: *J. Phys. Chem. Lett.* 14 (2023), pp. 7454–7460.

[194] Rebecca Hanscam and Eric Neuscamman. "Applying generalized variational principles to excited-state-specific complete active space self-consistent field theory". In: *J. Chem. Theory Comput.* 18.11 (2022), pp. 6608–6621.

[195] Taichi Yoneda et al. "Construction of sulfur-containing compounds with anti-cancer stem cell activity using thioacrolein derived from garlic based on nature-inspired scaffolds". In: *Org. Biomol. Chem.* 20.1 (2022), pp. 196–207.

[196] Noralhuda Ahmed Hamzah, Faehaa Azher Al-Mashhadane, and Suhad M Hamdoon. "Effects of Prepared Garlic (Allium Sativum) Extracts on C. Albicans and s. Aureus. Isolated from Oral Cavity". In: *J. Pharm. Negat.* (2022), pp. 794–801.

[197] Peter R Taylor. "Analytical MCSCF energy gradients: Treatment of symmetry and CASSCF applications to propadienone". In: *J. Comput. Chem.* 5.6 (1984), pp. 589–597.

[198] Richard Phillips Feynman. "Forces in molecules". In: *Phys. Rev.* 56.4 (1939), p. 340.

[199] James W Snyder et al. "An atomic orbital-based formulation of analytical gradients and nonadiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units". In: *J. Chem. Phys.* 143.15 (2015).

[200] James W Snyder et al. "A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units". In: *J. Chem. Phys.* 146.17 (2017).

[201] Nelson HF Beebe and Jan Linderberg. "Simplifications in the generation and transformation of two-electron integrals in molecular calculations". In: *Int. J. Quantum Chem.* 12.4 (1977), pp. 683–705.

[202] Henrik Koch, Alfredo Sánchez de Merás, and Thomas Bondo Pedersen. "Reduced scaling in electronic structure calculations using Cholesky decompositions". In: *J. Chem. Phys.* 118.21 (2003), pp. 9481–9484.

[203] Sarai D Folkestad, Eirik F Kjønstad, and Henrik Koch. "An efficient algorithm for Cholesky decomposition of electron repulsion integrals". In: *J. Chem. Phys.* 150.19 (2019).

[204] Lee-Ping Wang and Chenchen Song. "Geometry optimization made simple with translation and rotation coordinates". In: *J. Chem. Phys.* 144.21 (2016).

[205] Florian Weigend and Reinhart Ahlrichs. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy". In: *Phys. Chem. Chem. Phys.* 7 (2005), p. 3297. DOI: 10.1039/b508541a.

[206] H.-J. Werner et al. *MOLPRO, version 2019.2, a package of ab initio programs.* see https://www.molpro.net. Cardiff, UK, 2019.

[207] *Avogadro: an open-source molecular builder and visualization tool. Version 1.2.0.* http://avogadro.cc/.

[208] Marcus D Hanwell et al. "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform". In: *J. Cheminform.* 4.1 (2012), pp. 1–17.

[209] Kristopher T Jensen et al. "Modeling Electron Transfers Using Quasidiabatic Hartree–Fock States". In: *J. Chem. Theory Comput.* 14.9 (2018), pp. 4629–4639.