

A Parameterized Vibronic Spin-Orbit Coupling Model Protocol Suitable for Spectroscopy

by

Benny Jun Chen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Chemistry

Waterloo, Ontario, Canada, 2024

© Benny Jun Chen 2024

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Benny Chen was the sole author for Chapters 1, 3, 4, and 5, which were written under the supervision of Profs. Marcel Nooijen and Tao (Toby) Zeng. Chapter 2 was written in collaboration with Prof. Zeng, who contributed enormously to describing vibronic coupling theory. Feedback and revisions of the manuscript were given by Profs. Nooijen, Zeng, Hopkins. Chapter 3 had input provided by Dr. Neil Raymond on describing and designing the protocol. The research work presented in Chapter 3 and 4 was conducted jointly at both University of Waterloo and York University.

This thesis contains no work that has been published or submitted for publication.

Abstract

A diabaticization protocol constructing vibronic model Hamiltonians with inclusion of spin-orbit coupling was implemented in Python. This protocol has been extended to include spectroscopic applications. GAMESS package GMC-QDPT level of theory calculations carries out a proposed diabaticization scheme to automatically compute a grid of diabatic states in expansions of displaced nuclear coordinates. Diabatic potential energy surfaces can subsequently be constructed through fitting parameters. Generated vibronic models can be used for input to quantum dynamical simulation approaches such as MCTDH and VECC alike, where auto and cross-correlation functions can be obtained after propagation. Simulated gas-phase photoelectron spectra were reproduced for H_2O , NH_3 , and PH_3 , with excellent agreement to experimentally recorded spectra. Absorption spectra of transition metal trifluorides CoF_3 and RhF_3 , along with iron pentacarbonyl $\text{Fe}(\text{CO})_5$, were studied with a triple zeta polarized Sapporo basis set. Both CoF_3 and RhF_3 showed pronounced splitting originating from spin-orbit coupling effect, whereas $\text{Fe}(\text{CO})_5$ only displayed minimal change due to spin-orbit coupling limited to its truncated constant-order spectrum. It is predicted that the Jahn-Teller effect plays a more dominant role over spin-orbit coupling in our $\text{Fe}(\text{CO})_5$ model simulated spectra. Streamlining of the protocol has increased its accuracy and robustness, in the interest of supporting future vibronic spin-orbit coupling model research.

Acknowledgements

I am forever grateful to Prof. Nooijen for supervising me throughout the years. It was not clear to me at first, but the wisdom in connecting me with my future mentors and peers has been nothing short of magical. I also appreciate the extensive flexibility, facilitation, and support given in this project. Prof. Zeng has been an excellent and invaluable co-supervisor without whom this work would not be possible. His patience in advising me was simply remarkable. I will always cherish the immense kindness in offering me knowledge anytime I step into his office, no matter the question. The programming and computational aspects of the protocol would not be possible without the wizardry of Dr. Neil George Raymond, whose code is always breathtaking. I hope our fruitful collaborations never cease. My gratitude also go to Profs. Hopkins, Roy, Santoro, and Dieter Meyer. I am deeply appreciative of the continual love and support from my parents. Much thanks to Nooijen and Zeng group peers: Gabby, Eka, Songhao, Azh, and Coco for wonderful discussions and hangouts.

Table of Contents

List of Figures	ix
List of Tables	xii
List of Abbreviations	xiii
List of Listings	xiii
1 Introduction	1
1.1 Vibronic Models Background	1
1.2 Modern Vibronic Models - Timeline	3
1.3 Project Scope: Pivot to Zeng Protocol and Limitations of Single-Reference Linear Vibronic Models	5
1.4 Outline of Thesis	7
2 Theoretical Overview and Methodology	8
2.1 Electronic Structure Theory	8
2.1.1 Diabatic States	15
2.1.2 Our Diabatization Scheme	21
2.2 Building Wavefunctions	23
2.2.1 Multi-configuration Self Consistent Field (MCSCF)	23
2.2.2 CASSCF, RASSCF, CASPT, and GMC-QDPT	23
2.2.3 Active Space Design	25
2.3 Spin Orbit Coupling, Jahn-Teller Effect, and Pseudo-Jahn-Teller Effects	26
2.3.1 Spin Orbit Coupling	26
2.3.2 Jahn-Teller	28
2.3.3 Pseudo-Jahn-Teller	30

2.3.4	The Interplay between Spin-Orbit and (Pseudo-)Jahn-Teller	33
2.4	Spectroscopic Simulations	34
2.4.1	MCTDH	36
2.4.2	VECC	36
2.5	Computational Hardware and Software	37
2.5.1	GAMESS	37
2.5.2	nlogn, ComputeCanada (Computational Resources)	38
3	Diabatization Code Development	39
3.1	MCSCF Pipeline	40
3.2	Automated Machine Gun Diabatization Code	42
3.2.1	Step B: Frequencies and Reference Structure	45
3.2.2	Step C: Calculation at Reference Geometry (refG)	46
3.2.3	Step D: Submit Diabatization Displacement Calculations	47
3.2.4	Step E: Harvest Results for MCTDH/VECC	50
3.2.5	Step F: Fitting on n-th Order Models (Conditional)	54
3.3	Assorted Comments and Considerations	56
3.3.1	Useful Regular Expressions	56
3.3.2	Memory Mapping	56
3.3.3	Effective Vibronic Coupling Magnitude: Screening	57
3.3.4	Computational Roadblocks (<i>nlogn</i> and MCTDH)	57
4	Realistic Applications to Systems of Interest	60
4.1	Generating Spectra via MCTDH and VECC	61
4.1.1	MCTDH Propagations	61
4.1.2	SOC: Imaginary Changes Everything	63
4.1.3	VECC Propagations	64
4.2	Preliminary Systems	64
4.2.1	Water, H ₂ O	65
4.2.2	Ammonia, NH ₃	69
4.2.3	Phosphine, PH ₃	72
4.3	Complex Systems	74
4.3.1	Cobalt trifluoride, CoF ₃	74
4.3.2	Rhodium trifluoride, RhF ₃	78
4.3.3	Iron pentacarbonyl, Fe(CO) ₅	81

5 Conclusion	88
Bibliography	91
Appendices	102
A	103
A.1 MCSCF Pipeline	103
A.1.1 Step 1: Geometry optimization + frequency (Hessian)	103
A.1.2 Step 2: GMCPT	105
A.1.3 Step 3: Diabatic Molecular Orbitals (DMO)	106
A.1.4 Step 4: Reference Determinant	106
A.2 Formaldehyde, CH ₂ O	109

List of Figures

1.1	Stereotypical depiction of a conical intersection (CoIn) involving both adiabats and diabats commonly seen in vibronic models.	2
1.2	(a) Caffeine. Slightly matching peaks in the sub 240nm region observed, appears to need a linear red-shift for better agreement. (b) Titanium tetrachloride. The predictive accuracy of a smooth single peak is nowhere near reality, due to limitations of single reference TD-DFT methods applied to transition metal complexes. Santoro models also had no inclusion of SOC.	5
1.3	Original proposed research plan.	6
2.1	Depiction of a BOA Electronic Ground State Surface	14
2.2	Comparison of Adiabatic and Diabatic States	20
2.3	Quantized energy level diagram showing orbitals and electrons in a CAS(6,8) active space.	25
2.4	Comparison of D_{3h} and C_{3v} NH_3 orbitals, the D_{3h} HOMO a_2'' and undergoes Jahn-Teller distortion to stabilize as C_{3v} 's $2a_1$ lowers in energy. The C_{3v} NH_3 has bonding orbitals comparatively lower in energy than D_{3h} . Mixing interactions are highlighted in red.	32
2.5	Common Demonstration of the Franck-Condon Principle	35
3.1	GMC-QDPT vibronic coupling model scheme formulated by Zeng.	40
3.2	The six step process automated by the <code>dist</code> code.	43
3.3	Alternative <code>dist</code> flow chart, displaying the connections and dependencies between the various steps that occur when running the program.	44
3.4	Printing of complex matrices that are needed for SOC energies. The 17 rows and columns represents the electronic surfaces in a $\text{Fe}(\text{CO})_5$ calculation. Lower triangle values are complex conjugates of the upper triangle.	53
3.5	PES of D_{3h} NH_3 , Mode 7, +/- 10 Displacements, $q_{\text{size}} = 0.30$	55
3.6	PES of $\text{Fe}(\text{CO})_5$, 5 states, Mode 15, +/- 8 Displacements, $q_{\text{size}} = 0.05$	55
3.7	ComputeCanada usage breakdown during the 2024 academic year. The entirety of the second column CPU usage was purely dedicated to $\text{Fe}(\text{CO})_5$ diabaticization calculations.	58

4.1	Snapshots of a $\text{Fe}(\text{CO})_5$ 17 states MCTDH 50 fs propagation. The wavepacket density excites and then proceeds to relax.	62
4.2	Theoretical vibrationally-resolved photoelectron spectrum of H_2O	67
4.3	Theoretical vibrationally-resolved photoelectron spectrum of H_2O between 17–21 eV. A higher resolution photoelectron spectra with zoomed-in regions is available in Ref [79].	67
4.4	Theoretical vibrationally-resolved photoelectron spectrum of H_2O using MCTDH.	68
4.5	Theoretical vibrationally-resolved photoelectron spectrum of H_2O using VECC.	68
4.6	Theoretical vibrationally-resolved photoelectron spectrum of NH_3	70
4.7	Theoretical vibrationally-resolved photoelectron spectrum of NH_3 between 14–19 eV.	70
4.8	Theoretical vibrationally-resolved photoelectron spectrum of NH_3 using MCTDH.	71
4.9	Theoretical vibrationally-resolved photoelectron spectrum of NH_3 using VECC.	71
4.10	Theoretical vibrationally-resolved photoelectron spectrum of PH_3	73
4.11	Theoretical vibrationally-resolved photoelectron spectrum of PH_3	73
4.12	Theoretical vibrationally-resolved spectra of CoF_3 for three truncations of the vibronic model expansion. The upper panel includes only constant order contributions, the middle panel up to linear order, and the bottom panel up to quadratic order.	76
4.13	Theoretical vibrationally-resolved spectrum of CoF_3 using MCTDH.	77
4.14	Theoretical vibrationally-resolved spectrum of CoF_3 using VECC.	77
4.15	Theoretical vibrationally-resolved spectra of RhF_3 for three truncations of the Hamiltonian expansion. The upper panel includes only constant order contributions, the middle panel up to linear order, and the bottom panel up to quadratic order.	79
4.16	Theoretical vibrationally-resolved spectrum of RhF_3 using MCTDH.	80
4.17	Theoretical vibrationally-resolved spectrum of RhF_3 using VECC.	80
4.18	The $^1A'$ active space of 8o10e $\text{Fe}(\text{CO})_5$	82
4.19	3D view of the chosen 12 screened-modes (out of 27 possible modes) of $\text{Fe}(\text{CO})_5$ and their associated frequencies. The modes were filtered by means of an effective vibronic coupling magnitude calculation.	83
4.20	SOC eigenvalues spectrum for $\text{Fe}(\text{CO})_5$ <i>refG</i> GMC-QDPT energies.	84
4.21	Theoretical vibrationally-resolved spectrum of $\text{Fe}(\text{CO})_5$. Lower energy peak not shown. MCTDH comparison of SOC effect. Propagated for 180 fs and resolved with $\tau = 40$	86
4.22	Theoretical vibrationally-resolved spectrum of $\text{Fe}(\text{CO})_5$. Lower energy peak not shown. VECC comparison of SOC effect. Propagated for 180 fs and resolved with $\tau = 40$	86

4.23	Theoretical vibrationally-resolved linear order spectrum of $\text{Fe}(\text{CO})_5$. Propagated for 180 fs and resolved with $\tau = 40$	87
4.24	Theoretical vibrationally-resolved quadratic order spectrum of $\text{Fe}(\text{CO})_5$. Propagated for 180 fs and resolved with $\tau = 40$	87
A.1	RhF3	103
A.2	Theoretical vibrationally-resolved photoelectron spectrum of CH_2O	110
A.3	Theoretical vibrationally-resolved photoelectron spectrum of CH_2O	110
A.4	Composite of theoretical vibrationally-resolved spectrum of CoF_3	111
A.5	Composite of theoretical vibrationally-resolved spectrum of RhF_3	112
A.6	Theoretical vibrationally-resolved spectrum composite of $\text{Fe}(\text{CO})_5$. Propagated for 180fs and resolved with $\tau = 40$	113
A.7	Theoretical vibrationally-resolved spectrum of $\text{Fe}(\text{CO})_5$ between 6–12 eV. Propagated for 180fs and resolved with $\tau = 40$. This region is quite sensitive to the contribution from varying parameters such as basis function size and restricted modes.	113

List of Tables

2.1	Character Table and IRREPs for the D_{3h} Point Group. Since the first order SOC requires a spatially degenerate term symbol with non-zero spin, we focused on ${}^2E$ -type term symbols in our study [39, 58].	28
2.2	Comparison of the MCTDH and VECC approaches, highlighting their respective advantages and capabilities. Both methods can use the same vibronic Hamiltonian input. The polynomial $O(n^2)$ scaling is in terms of n normal modes. VECC can calculate thermal properties and is configured to only use simple on/off flags.	37
4.1	H ₂ O model computed parameters. All numerical values are in eV. The model includes constant, linear, bilinear, and up to quadratic order vibronic coupling terms.	66
A.1	Step 1 Information	104
A.2	Step 2 Information	105
A.3	Step 3 Information	107
A.4	Step 4 Information	107

List of Abbreviations

BOA - Born-Oppenheimer Approximation
CASSCF - Complete Active Space SCF
CI - Configuration Interaction
CMO - Canonical Molecular Orbitals
dist - Diabatization Python Program 'dist'
DMO - Diabatic Molecular Orbitals
DSOME - Diabatic Spin-Orbit Matrix Elements
GAMESS - General Atomic and Molecular Electronic Structure System
GMC-QDPT - Generalized Multiconfiguration Quasi-Degenerate Perturbation Theory
IRREP - Irreducible Representations
(p)JT - (pseudo-) Jahn-Teller
MCSCF - Multiconfiguration Self Consistent Field
MCTDH - Multiconfiguration Time-Dependent Hartree
ME - Matrix Elements
MR - Multireference
NACs - Non-Adiabatic Couplings
PES - Potential energy surface
PT - Perturbation Theory
REFDET - Reference Determinants
refG - Reference Geometry
SCF - Self Consistent Field, Hartree-Fock (HF) method
SD - Slater Determinant
SOC - Spin Orbit Coupling
TD-DFT - Time-Dependent Density Functional Theory
TDM - Transition Dipole Moment
VECC - Vibrational Electronic Coupled Cluster

List of Listings

3.1	Optimized geometry of $\text{Fe}(\text{CO})_5$ coordinates at equilibrium geometry in Å. . .	45
3.2	Optimized frequencies of $\text{Fe}(\text{CO})_5$	46
3.3	TDM block as seen in the $\text{Fe}(\text{CO})_5$ refG calculation.	47
3.4	Precomputing linear, quadratic, and bilinear displacements in dist	47
3.5	-x1q31_-x1q27 distorted coordinates (negative 31st and 27th mode linear order).	48
3.6	$\text{Fe}(\text{CO})_5$ 17 states rotation matrix U calculated within GAMESS.	49
3.7	MCTDH operator file format constructed with standard ingredients in a vibronic model.	50
3.8	Differentiation of linear vibronic coupling values extraction in the dist program.	51
3.9	Diabatic matrix elements example found in GAMESS.	51
3.10	Evaluating linear coupling term and conversion of units. Not the full code implementation in dist , demonstration purposes only.	51
3.11	Evaluating quadratic coupling term and conversion of units. Not the full code implementation in dist , demonstration purposes only.	52
3.12	Evaluating bilinear coupling term and conversion of units. Not the full code implementation in dist , demonstration purposes only.	52
3.13	DSOME extraction involves summing the one and two electron values for real and imaginary components.	53
4.1	Pedagogical example of water active space parameters in GAMESS input file.	65
A.1	Example D_{3h} RhF_3 Input Geometry	104
A.2	Additional flags for Step 1.	104
A.3	Additional flags for Step 2.	105
A.4	Additional flags for Step 3.	106
A.5	Additional flags for Step 4.	107
A.6	Additional flags for Step 5.	108

Chapter 1

Introduction

1.1 Vibronic Models Background

Since the breakthrough of quantum physics, it was claimed that the whole of chemistry was in essence solved. The Schrödinger equation had it all, as long as you were able to solve it. Despite the rapid growth in computing, for the majority of meaningful and sizeable chemical systems, the calculation of its exact solutions remains intractable. Therefore, modern quantum chemistry relies on clever mathematical approximations to calculate energies, sufficient for predictive and chemical accuracy. In 1927, Born and Oppenheimer proposed the Born-Oppenheimer Approximation (BOA), a method that separates the nuclear and electronic regimes and allows one to calculate the energy of adiabatic potential energy surfaces (PES). The transformed electronic Hamiltonian was more manageable. However, many realistic chemical processes are innately non-adiabatic, and have electronic states that are not well separated and are nearly degenerate. Prime examples where non-adiabatic effects show an important role are the photo-isomerization of retinal and transition metal complex spectra [1, 2, 3, 4]. Vibrational-electronic (*vib-ronic*) coupling models aim to simulate the mechanics revolving around non-adiabatic dynamics. Maturation in both post-Hartree-Fock and computing methods have since enabled theoretical chemists to explore sophisticated non-adiabatic phenomena such as conical intersections and intersystem crossings [5, 6].

Research in non-adiabatic chemistry has since gained momentum in recent decades leading to the construction of vibronic models that can go beyond the BOA and can capture a proper non-adiabatic picture. The use of vibronic models in a diabatic basis is theorized to be one viable approach to non-adiabatic treatments. This procedure is referred to as a diabaticization. Diabatic states (diabats) have the advantage of diminishing the tricky non-adiabatic couplings (NACs) that arise from the BOA adiabatic states (adiabats), and maintaining smooth character along the PES. To solve for dynamics, many proceed to use the Multiconfiguration Time-dependent Hartree (MCTDH) algorithm. A long term goal of the Nooijen group is to formulate novel alternative operator techniques like Vibrational Electronic Coupled Cluster (VECC) to calculate vibrationally-resolved electronic spectra and statistical mechanics based on a vibronic model framework [7, 8, 9].

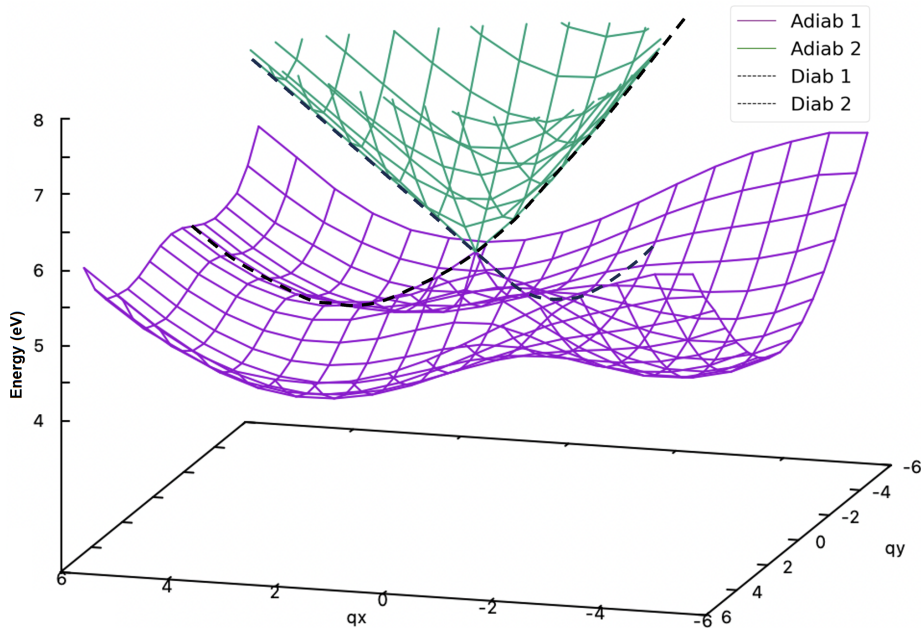


Figure 1.1: Stereotypical depiction of a conical intersection (CoIn) involving both adiabats and diabats commonly seen in vibronic models.

The central goal in this thesis is to produce a parameterized diabaticization protocol that can supply MCTDH and VECC with vibronic model Hamiltonians suitable for a wide range of applications such as spin-orbit coupling (SOC) effects and spectroscopy. For the past two years at the Nooijen and Zeng group, we have jointly endeavoured to pursue this goal.

1.2 Modern Vibronic Models - Timeline

The pioneering work of Cederbaum, Köppel, and Domcke gave rise to the original idea of vibronic models in their 1981 seminal paper [10]. Determination of a vibronic model Hamiltonian that simultaneously coupled vibrational modes with an arbitrary set of electronic states was given in an *ab initio* procedure. Their study proceeded to simulate spectroscopy and demonstrated the need for the inclusion of vibronic coupling effects to match with findings in experimental spectra. Among others, Neugebauer and Nooijen found this approach useful and further developed a diabaticization scheme suitable for Time-dependent Density Functional Theory (TD-DFT) absorption spectrum [11, 12]. Nooijen, Chang, and Hazra also developed vibronic models based on Equation of Motion Coupled Cluster approaches, in particular for ionized states and to simulate photoelectron spectra [13, 14, 15]. Vibronic models were used to model multiple potential energy surfaces in a limited region of configuration space, akin to a harmonic oscillator approximation applied to a single surface. Moreover, this established a formalized, standard method for routinely extracting vibronic model Hamiltonians from quantum chemistry calculations. Several Nooijen group members have since iterated on these methods [16, 17]. Nooijen’s work inspired Santoro and coworkers, who in recent years recasted this scheme into a single-reference linear order Gaussian-MCTDH TD-DFT diabaticization scheme purposed for studying excited state dynamics of hexahelicene [18, 19, 20]. However, the electronic structure of hexahelicene is relatively straightforward, as its ground and excited states are well described by standard DFT and TD-DFT approaches. Ongoing frontier research by Domcke and Mondal aims to apply vibronic models for transition metal complex spectra, a notoriously difficult task due to the dilemma of low-lying nearly-degenerate electronic states. It is difficult to resolve the states and PES character for such systems. Therefore, a more ideal approach is to pivot towards a diabaticization scheme that is multi-reference in nature, and can capture relativistic effects like spin-orbit coupling.

Such a spin-orbit-coupling-capable diabaticization protocol was devised by Zeng [21], which has been successfully used in subsequent studies of the Zeng group [22, 23, 24, 25, 26, 27].

The foundation of the Zeng SOC protocol derives from previous work done by Atchity and Rudenberg on configurational uniformity as well as model space diabaticization (MSD) by Li et. al [28, 29, 30], and specifically uses GAMESS-US package Generalized Multi-Configuration Quasidegenerate Perturbation Theory (GMC-QDPT) calculations [31, 32, 33, 34]. Diabatic states are directly calculated using Nakamura and Truhlar’s version of configurational uniformity in GAMESS, whereby diabatic molecular orbitals (DMOs) are first prepared, and then computation of diabatic state functions follows [35, 36]. Calculations involved here are typed as ‘multi-state multi-reference perturbation theory’ calculations in a given active space reference as implemented by Nakano. GMC-QDPT addresses static correlation using Multi-configuration Self Consistent Field (MCSCF) (specifically Complete Active Space SCF (CASSCF)) wavefunctions, and then dynamical correlation is captured by perturbation theory (PT). At this point, Zeng assembles all these pieces and opens up the possibility of adding SOC terms to the model Hamiltonian. The basis of the vibronic models in this thesis will use the diabatic energies and couplings yielded by this process.

The use of vibronic models, including SOC, to simulate spectroscopy has been used before. Siebrand and Zgierski fitted vibronic SOC interactions in their 1979 pyrazine study [37]. On transition metals, Nakano et al. has scrutinized the behaviour of FeCO^+ in a wavefunction study [38]. First author Mondal has in particular discussed $E \otimes e$ Jahn-Teller (JT) and SOC effects in transition-metal fluorides [39, 40]. Analysis included generation of trigonal planar D_{3h} symmetry vibronic model Hamiltonians and spectra. Previous applications of the Zeng SOC diabaticization protocol were towards the construction of vibronic models for pnictogen hydride cations with trigonal symmetry [21] and the simulations of singlet fission dynamics [22, 23, 24, 25, 26, 27]. No spectroscopy components were included. So far, this context and state of research converges on a case, to which the project’s diabaticization protocol should be applied: a large D_{3h} iron-carbonyl complex with low-lying degenerate electronic states. The candidate molecule is $\text{Fe}(\text{CO})_5$. If successful, the appealing challenge of $\text{Fe}(\text{CO})_5$ should exhibit demonstrable JT, SOC, and photochemical effects as purported in the literature [41].

1.3 Project Scope: Pivot to Zeng Protocol and Limitations of Single-Reference Linear Vibronic Models

In early preliminary work during an undergraduate research project and this thesis, I gained experience with the Santoro G-MCTDH models. As previously mentioned, transition metal complexes are appealing vibronic systems to probe, and SOC is needed for proper characterization. Substances like TiO_2 and TiCl_4 were familiar to me due to working at a paint and coatings factory. Predictably, Santoro models did not suffice for transition metal complexes. As shown in Figure 1.2, the vibronic models were found to not resolve any vibronic fine structure in metallic spectra, and for cases such as caffeine, the spectrum was still inaccurate despite meticulous scrutiny of the model. Santoro has likewise acknowledged computational study inaccuracies due to complications in finding correct parameters for calculations [19].

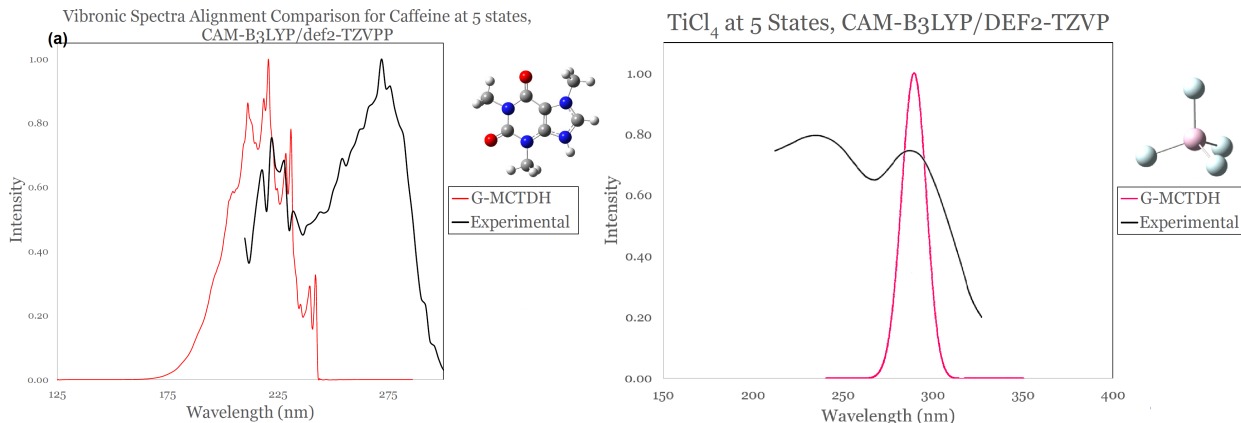


Figure 1.2: (a) Caffeine. Slightly matching peaks in the sub 240nm region observed, appears to need a linear red-shift for better agreement. (b) Titanium tetrachloride. The predictive accuracy of a smooth single peak is nowhere near reality, due to limitations of single reference TD-DFT methods applied to transition metal complexes. Santoro models also had no inclusion of SOC.

To capitalize on all the work presented so far, I have created a novel automated diabaticization code in Python to be incorporated into the Zeng protocol. This is my main contribution. Such limitations of the Santoro method resulted in a pivoting towards the Zeng protocol, in order to enhance the capabilities of vibronic models available to VECC. Previously, only the top half path shown in Figure 1.3 (Santoro-MCTDH-Spectra) was explored. While VECC could perform excited state simulations like MCTDH, its thermochemistry module should

align with Zeng models for theoretical coherence [9]. To declare our task at hand: we shall construct a composite approach where the Zeng protocol is first used to manufacture vibronic model Hamiltonians to be plugged-in as input, and the latter half of non-adiabatic simulation will be taken care of by either MCTDH or VECC. A Zeng-VECC-Spectra pathway possibility. Once the Zeng protocol can be successfully connected to both MCTDH and VECC, then it will allow for a variety of models and dataset to be used for practical comparison of the gold standard MCTDH method with the new VECC method.

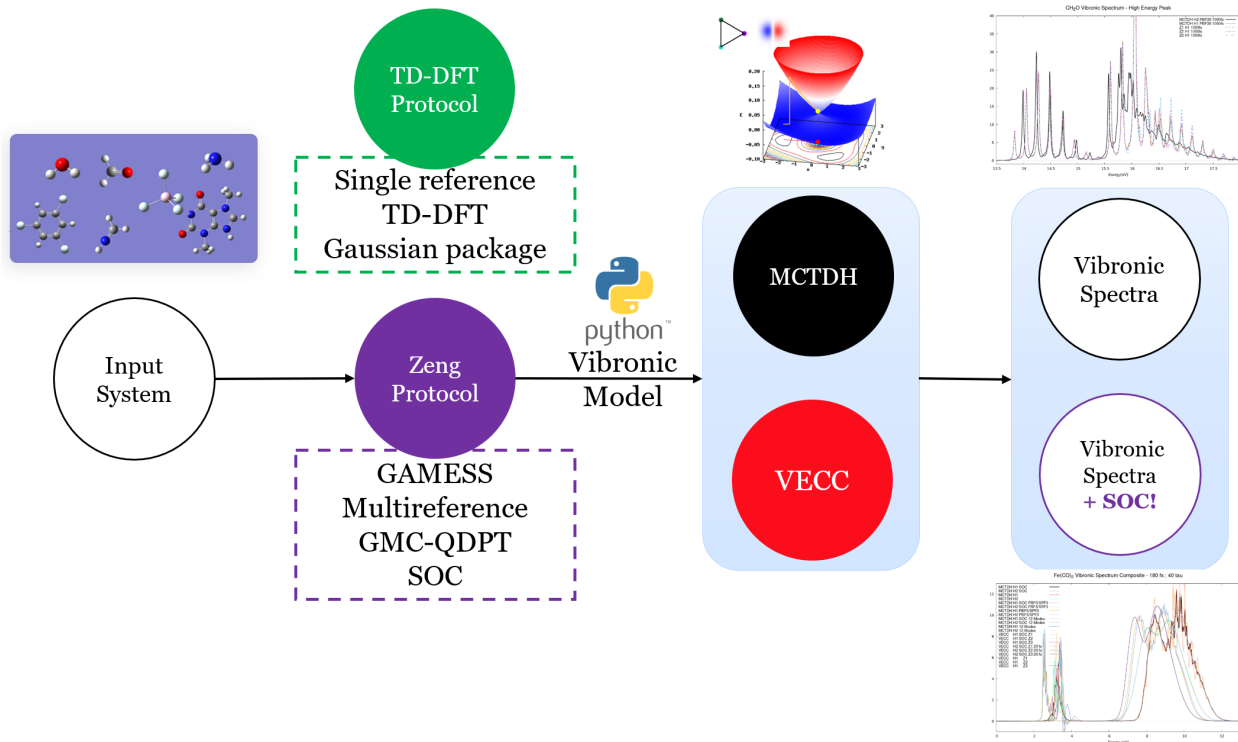


Figure 1.3: Original proposed research plan.

Formerly, the Zeng diabatisation protocol was not automated. Each calculation in the series needed for full diabatisation was handcrafted. In early 2023, Zeng completed the first step towards accomplishing the task: a prototype 'machine gun' Bash code that automates diabatisation. This version could effectively produce up to second-order Zeng non-SOC models, and was implemented to be extracted from GAMESS calculations. Neil Raymond and I have iterated on the prototype and designed a seamless operating protocol. The code was first translated into Python, and then received a wide range of upgrades to ensure that it is SOC-compatible and readily plugged into MCTDH/VECC for spectroscopy.

1.4 Outline of Thesis

The rest of the thesis is organized as follows. In Chapter 2, fundamental quantum chemistry concepts and relevant mathematics knowledge are presented, with an emphasis on diabatisation. Details on active space, SOC/JT effects, and quantum dynamics simulation approaches are included. The specific diabatisation scheme and methodology used to conduct the calculations are contained within Chapter 3. The main purpose of this third chapter is devoted to the full elucidation of the Python code to automate the Zeng protocol. Various roadblocks and design decisions that went into the development are also discussed. Applications of the protocol to practical small organic molecules and realistic transition metal systems are shown in Chapter 4. The simulated spectroscopic results are compared, analyzed, and evaluated. In Chapter 5, the last chapter, a future outlook and the conclusions are presented.

Chapter 2

Theoretical Overview and Methodology

2.1 Electronic Structure Theory

In an ideal world, the Schrödinger equation would be readily solved by computers and the exact analytical solution for energy eigenvalues would be known. However, in reality this is not computationally viable for even simple second row atoms—let alone molecular systems.

We will break down the mathematics that underlie the BOA, vibronic Hamiltonians, and how one reaches the all important unitary transformation matrix essential to diabaticization.

The derivation will follow the lead and repeat the most salient parts of Refs. [21, 42, 43].

Let the full molecular Schrödinger’s Equation be defined as:

$$\hat{H}\Psi_n(\vec{r}, \vec{R}) = E_n\Psi_n(\vec{r}, \vec{R}) \quad (2.1)$$

The full molecular Hamiltonian:

$$\hat{H} = -\sum_{\alpha}^M \frac{1}{2M_{\alpha}} \nabla_{\alpha} \cdot \nabla_{\alpha} - \sum_i^N \frac{1}{2} \nabla_i \cdot \nabla_i + \sum_{i,j>i}^N \frac{1}{r_{ij}} + \sum_{i,\alpha}^{N,M} \frac{-Z_A}{r_{i\alpha}} + \sum_{\alpha,\beta>\alpha}^M \frac{Z_A Z_B}{R_{\alpha\beta}} \quad (2.2)$$

Which can concisely be represented by kinetic and potential energy operators $\hat{T}$ and $\hat{V}$:

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne} + \hat{V}_{NN} \quad (2.3)$$

Note the subscripts e, N are for electronic and nuclear correspondingly. We can decouple the kinetic energy of the nuclei and electronic Hamiltonian into split parts.

$$\hat{H} = \hat{T}_N + \hat{H}_{elec} \quad (2.4)$$

Where the electronic Hamiltonian is given by:

$$\hat{H}_{elec} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne} + \hat{V}_{NN} \quad (2.5)$$

One may consider the $\hat{T}_N$ and $\hat{H}_{elec}$ separately. Fixed nuclei wavefunctions are obtained from this electronic structure problem. In this manner, one first solves the Schrödinger equation for $\hat{H}_{elec}$ at each $\vec{R}$:

$$\hat{H}_{elec} \left| \phi_{\lambda}^{ad}(\vec{R}) \right\rangle = E_{\lambda}^{ad}(\vec{R}) \left| \phi_{\lambda}^{ad}(\vec{R}) \right\rangle. \quad (2.6)$$

The electronic eigenstates $\left\{ \left| \phi_{\lambda}^{ad}(\vec{R}) \right\rangle \right\}$ and the eigenenergies $\left\{ E_{\lambda}^{ad}(\vec{R}) \right\}$ parametrically depend on the nuclear configuration $\vec{R}$, which itself is a parameter for adiabatic states. $\vec{r}$ denotes the electron coordinates. Those electronic eigenstates are called adiabatic states (abbreviated as adiabats) and their energies are called adiabatic energies. That is why they are decorated by the superscript “ ad ”. $\left\{ \left| \phi_{\lambda}^{ad}(\vec{R}) \right\rangle \right\}$ forms a complete set for the electronic degrees of freedom and can be used to expand any vibronic states without loss of generality:

$$\Psi(\vec{r}, \vec{R}) = \sum_{\lambda} \phi_{\lambda}^{ad}(\vec{r}; \vec{R}) \chi_{\lambda}^{ad}(\vec{R}) \quad (2.7)$$

Here, $\left\{ \left| \chi_{\lambda}^{ad} \right\rangle \right\}$ is a set of vibrational states, or more generally speaking, states of nuclear degrees of freedom, that are associated to the adiabatic electronic states in a one-to-one

manner. They can be understood as vibrational wave packets evolving on individual electronic adiabatic potential energy surfaces.

This expansion using adiabatic electronic states is equivalent to resolving the total Hamiltonian $\hat{H}$:

$$\begin{aligned}
\hat{H} &= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \hat{H} |\phi_\mu^{ad}\rangle \langle \phi_\mu^{ad}| \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \hat{T}_N + \hat{H}_{elec} |\phi_\mu^{ad}\rangle \langle \phi_\mu^{ad}| \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \left(\langle \phi_\lambda^{ad}| \hat{T}_N |\phi_\mu^{ad}\rangle + \delta_{\lambda\mu} E_\mu^{ad}(\vec{R}) \right) \langle \phi_\mu^{ad}| \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \hat{T}_N |\phi_\mu^{ad}\rangle \langle \phi_\mu^{ad}| + \sum_\lambda E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \quad (2.8)
\end{aligned}$$

Now, we write the nuclear kinetic operator as

$$\hat{T}_N = \frac{\hat{\vec{P}} \cdot \hat{\vec{P}}}{2} \quad (2.9)$$

where $\hat{\vec{P}}$ is the mass-weighted nuclear momentum operator for all nuclear degrees of freedom.

Then the Hamiltonian resolved using adiabatic states reads:

$$\hat{H} = \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \frac{\hat{\vec{P}} \cdot \hat{\vec{P}}}{2} |\phi_\mu^{ad}\rangle \langle \phi_\mu^{ad}| + \sum_\lambda E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \quad (2.10)$$

Considering that $\{\phi^{ad}\}$ parametrically depends on $\vec{R}$ [43, 44], and $\hat{\vec{P}} = -i\vec{\nabla}_R$ in the $\vec{R}$ -position representation, the result of $\hat{\vec{P}} \cdot \hat{\vec{P}} |\phi_\mu^{ad}\rangle$, applying the product rule of differentiation twice for the back-to-back operations of $\hat{\vec{P}}$, reads:

$$\begin{aligned}
\hat{\vec{P}} \cdot \hat{\vec{P}} |\phi_\mu^{ad}\rangle &= \hat{\vec{P}} \cdot \left| \hat{\vec{P}} \phi_\mu^{ad} \right\rangle + \hat{\vec{P}} |\phi_\mu^{ad}\rangle \cdot \hat{\vec{P}} \\
&= \left| \hat{P}^2 \phi_\mu^{ad} \right\rangle + \left| \hat{\vec{P}} \phi_\mu^{ad} \right\rangle \cdot \hat{\vec{P}} + \left| \hat{\vec{P}} \phi_\mu^{ad} \right\rangle \cdot \hat{\vec{P}} + |\phi_\mu^{ad}\rangle \hat{P}^2 \\
&= \left| \hat{P}^2 \phi_\mu^{ad} \right\rangle + 2 \left| \hat{\vec{P}} \phi_\mu^{ad} \right\rangle \cdot \hat{\vec{P}} + |\phi_\mu^{ad}\rangle \hat{P}^2 \quad (2.11)
\end{aligned}$$

Here, $|\hat{P}^2\phi_\mu^{ad}\rangle$ and $|\hat{\vec{P}}\phi_\mu^{ad}\rangle$ means the resultant electronic states after acting $\hat{P}^2$ or $\hat{\vec{P}}$ on the adiabatic state $|\phi_\mu^{ad}\rangle$. Substituting Eq. 2.11 back to Eq. 2.10, we have

$$\begin{aligned}
\hat{H} &= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \left(\frac{1}{2} \langle \phi_\lambda^{ad} | \hat{P}^2 \phi_\mu^{ad} \rangle + \langle \phi_\lambda^{ad} | \hat{\vec{P}} \phi_\mu^{ad} \rangle \cdot \hat{\vec{P}} + \langle \phi_\lambda^{ad} | \phi_\mu^{ad} \rangle \hat{T}_N \right) \langle \phi_\mu^{ad} | \\
&\quad + \sum_{\lambda} E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \left(\frac{1}{2} \langle \phi_\lambda^{ad} | \hat{P}^2 \phi_\mu^{ad} \rangle + \langle \phi_\lambda^{ad} | \hat{\vec{P}} \phi_\mu^{ad} \rangle \cdot \hat{\vec{P}} + \delta_{\lambda\mu} \hat{T}_N \right) \langle \phi_\mu^{ad} | \\
&\quad + \sum_{\lambda} E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \left(\frac{1}{2} \langle \phi_\lambda^{ad} | \hat{P}^2 \phi_\mu^{ad} \rangle + \langle \phi_\lambda^{ad} | \hat{\vec{P}} \phi_\mu^{ad} \rangle \cdot \hat{\vec{P}} \right) \langle \phi_\mu^{ad} | \\
&\quad + \left[\hat{T}_N + \sum_{\lambda} E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \right] \\
&= \sum_{\lambda,\mu} |\phi_\lambda^{ad}\rangle \left(-\frac{1}{2} G_{\lambda\mu}(\vec{R}) - i \vec{F}_{\lambda\mu}(\vec{R}) \cdot \hat{\vec{P}} \right) \langle \phi_\mu^{ad} | + \left[\hat{T}_N + \sum_{\lambda} E_\lambda^{ad}(\vec{R}) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \right] \\
&= \sum_{\lambda \neq \mu} |\phi_\lambda^{ad}\rangle \left(-\frac{1}{2} G_{\lambda\mu}(\vec{R}) - i \vec{F}_{\lambda\mu}(\vec{R}) \cdot \hat{\vec{P}} \right) \langle \phi_\mu^{ad} | \\
&\quad + \left[\hat{T}_N + \sum_{\lambda} \left(E_\lambda^{ad}(\vec{R}) - \frac{1}{2} G_{\lambda\lambda}(\vec{R}) \right) |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}| \right]
\end{aligned} \tag{2.12}$$

where the definitions of the $G_{\lambda\mu}$ matrix element and the $\vec{F}_{\lambda\mu}$ vectorial matrix element are self-evident; they are explicitly written as functions of $\vec{R}$, and the diagonal $\vec{F}_{\lambda\lambda} = 0$ due to the orthonormality (must be) and the real-valuedness (assumed in the present work) of $\{\phi^{ad}\}$. The real-valuedness is assumed since all non-spin-orbit electronic structure calculations in this thesis generate real-valued adiabats. So, it is not an assumption, but therefore a setting. One more detail to mention, the $\underline{\vec{F}}$ matrix is an anti-Hermitian matrix, since it is multiplied by the anti-Hermitian $-i$ to make a Hermitian operator.

In this adiabats-resolved $\hat{H}$, the terms in brackets $[\dots]$ govern the evolution of nuclear wave packets on individual adiabatic PESs with the diagonal Born-Oppenheimer corrections (DBOC) $\frac{1}{2}G_{\lambda\lambda}(\vec{R})$ [45]. The evolution of nuclear wave packet on one adiabatic PES is a typical Born-Oppenheimer approximation. The $\sum_{\lambda \neq \mu}$ terms govern the transitions of nuclear wave packets from one adiabat to another. Since those terms, $G_{\lambda\mu}$ and $\vec{F}_{\lambda\mu}$, breakdown the BO approximation, they are called non-adiabatic coupling matrix elements (NACME) [45, 46]. More concisely, people only call $\vec{F}_{\lambda\mu}$ the non-adiabatic coupling matrix elements. This is because the $\underline{\underline{G}}$ matrix can be written in terms of the $\underline{\underline{F}}$ vectorial matrix:

$$\begin{aligned}
(\vec{\nabla} \cdot \vec{F}_{\lambda\mu}) &= (\vec{\nabla} \cdot \langle \phi_\lambda^{ad} | \vec{\nabla} \phi_\mu^{ad} \rangle) \\
&= \langle \vec{\nabla} \phi_\lambda^{ad} | \cdot | \vec{\nabla} \phi_\mu^{ad} \rangle + \underbrace{\langle \phi_\lambda^{ad} | \nabla^2 \phi_\mu^{ad} \rangle}_{G_{\lambda\mu}} \\
&= \langle \vec{\nabla} \phi_\lambda^{ad} | \underbrace{\sum_\kappa |\phi_\kappa^{ad}\rangle \langle \phi_\kappa^{ad}| \cdot | \vec{\nabla} \phi_\mu^{ad} \rangle}_{\hat{I} = \sum_\kappa |\phi_\kappa^{ad}\rangle \langle \phi_\kappa^{ad}|} + G_{\lambda\mu} \\
&= \sum_\kappa \langle \vec{\nabla} \phi_\lambda^{ad} | \phi_\kappa^{ad} \rangle \cdot \langle \phi_\kappa^{ad} | \vec{\nabla} \phi_\mu^{ad} \rangle + G_{\lambda\mu} \\
&= \sum_\kappa \langle \phi_\kappa^{ad} | \vec{\nabla} \phi_\lambda^{ad} \rangle \cdot \langle \phi_\kappa^{ad} | \vec{\nabla} \phi_\mu^{ad} \rangle + G_{\lambda\mu} \\
&= \sum_\kappa \vec{F}_{\kappa\lambda} \cdot \vec{F}_{\kappa\mu} + G_{\lambda\mu} \\
&= - \sum_\kappa \vec{F}_{\lambda\kappa} \cdot \vec{F}_{\kappa\mu} + G_{\lambda\mu} \\
&= - (\underline{\underline{F}} \cdot \underline{\underline{F}})_{\lambda\mu} + G_{\lambda\mu} \\
G_{\lambda\mu} &= (\vec{\nabla} \cdot \vec{F}_{\lambda\mu}) + (\underline{\underline{F}} \cdot \underline{\underline{F}})_{\lambda\mu} \\
\underline{\underline{G}} &= (\vec{\nabla} \cdot \underline{\underline{F}}) + \underline{\underline{F}} \cdot \underline{\underline{F}}
\end{aligned} \tag{2.13}$$

So, all the non-adiabatic coupling is contained in the $\underline{\underline{F}}$ vectorial matrix. Starting from $\nabla \langle \phi_\lambda^{ad} | \hat{H}_{elec} | \phi_\mu^{ad} \rangle = 0$ (due to orthonormality of the two different adiabats), we can derive:

$$\vec{F}_{\lambda\mu} = \frac{\langle \phi_\lambda^{ad} | \vec{\nabla} \hat{H}_{elec} | \phi_\mu^{ad} \rangle}{E_\mu - E_\lambda} \quad (2.14)$$

where we can see that the magnitude of the vectorial NACME increases inversely proportional to the reciprocal of the adiabatic energy gap, and becomes singular when the adiabats approach degeneracy, i.e., non-adiabatic coupling (i.e., the breakdown of BOA) becomes significant for degenerate or pseudo-degenerate adiabats. It is worth mentioning that while the diagonal $\vec{F}_{\lambda\lambda} = 0$ always, the diagonal $G_{\lambda\lambda}$ gives a finite correction to the adiabatic potential except when $|\phi_\lambda^{ad}\rangle$ becomes degenerate with another (other) state(s), in which situation $G_{\lambda\lambda} = -\infty$. This is because:

$$\begin{aligned} G_{\lambda\lambda} &= \left(\vec{\nabla} \cdot \vec{F}_{\lambda\lambda} \right) + \left(\vec{F} \cdot \vec{F} \right)_{\lambda\lambda} \\ &= \left(\vec{\nabla} \cdot \vec{0} \right) + \sum_{\kappa \neq \lambda} \vec{F}_{\lambda\kappa} \cdot \vec{F}_{\kappa\lambda} \\ &= - \sum_{\kappa \neq \lambda} \vec{F}_{\lambda\kappa} \cdot \vec{F}_{\lambda\kappa} \\ &= - \sum_{\kappa \neq \lambda} \left| \vec{F}_{\lambda\kappa} \right|^2 \end{aligned} \quad (2.15)$$

When $|\phi_\lambda^{ad}\rangle$ comes close in energy to, say $|\phi_{\lambda+1}^{ad}\rangle$, the summation will be dominated by the singular $-\left| \vec{F}_{\lambda\lambda+1} \right|^2$ term, i.e., $G_{\lambda\lambda} = -\infty$, and the $-\frac{1}{2}G_{\lambda\lambda}$ DBOC correction becomes $+\infty$. This ∞ potential term determines that the wave packet on E_λ^{ad} cannot reach a nuclear configuration of degeneracy between $|\phi_\lambda\rangle$ and $|\phi_{\lambda+1}\rangle$ (or $|\phi_{\lambda-1}\rangle$) without hopping to $E_{\lambda+1}^{ad}$ (or $E_{\lambda-1}^{ad}$), i.e., a non-adiabatic transition has to occur.

This adiabatic representation features the following advantages: (1) $\{E^{ad}\}$ have the meaning of electronic energy levels as they are eigenvalues of $\hat{H}_{elec}$; (2) it matches the conventional photochemistry picture of non-adiabatic transition between electronic energy levels. For instance, the conventional Jablonski diagrams are drawn with respect to adiabats; (3) conven-

tional quantum chemistry calculations give adiabatic electronic states and their energies; (4) non-adiabatic transitions occur at specific locations in $\vec{R}$ with pseudo-degeneracy or degeneracy of adiabats; (5) the adiabatic PESs $\{E^{ad}\}$ are obtained from calculation at individual $\vec{R}$ points without the need of information at other $\vec{R}$ points as shown in Figure 2.1.

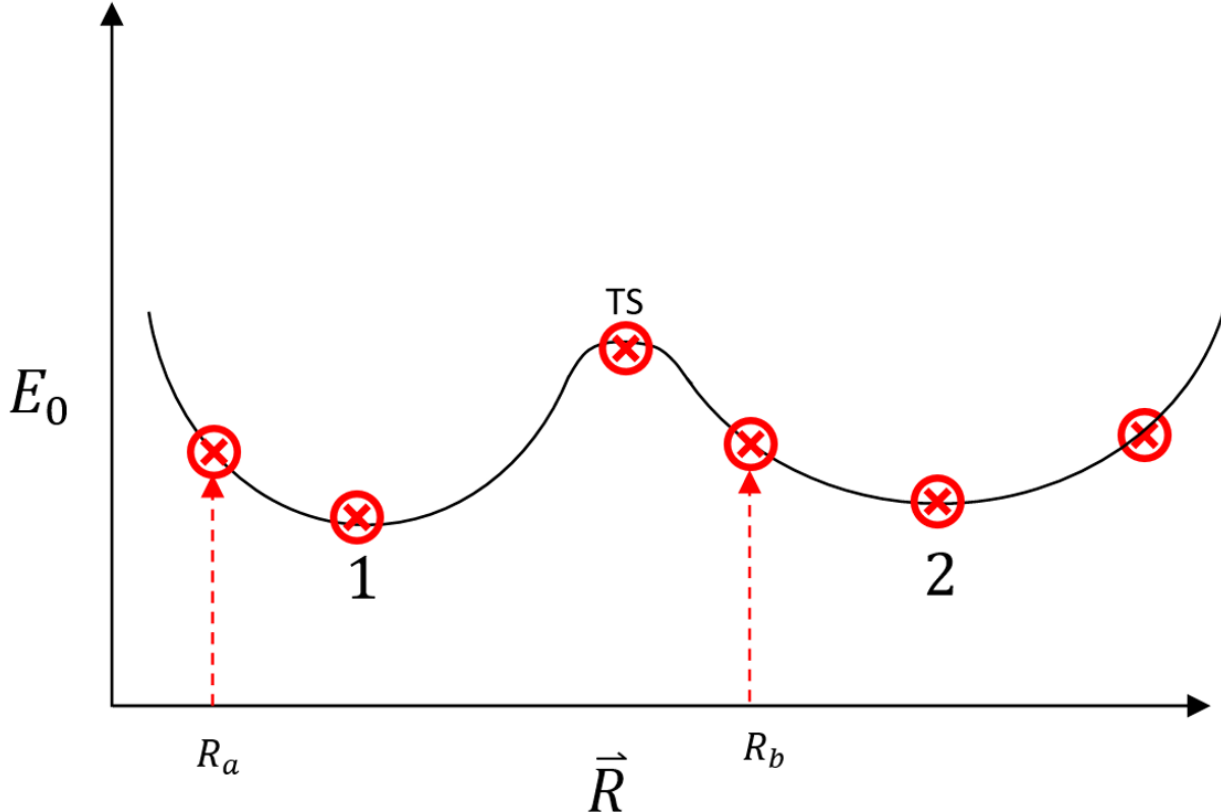


Figure 2.1: A single electronic ground state surface within the BOA leads to sampling and subsequently calculating points at fixed nuclear configurations $(\vec{R}_a, \dots, \vec{R}_b, \dots)$.

The adiabatic representation, however, features the following disadvantages. (1) Singular $\vec{F}_{\lambda\mu}$ and $G_{\lambda\lambda}$ when adiabatic electronic states are degenerate; (2) $\vec{F}_{\lambda\mu}$ may not be easy to calculate as explained in equation 2.14; (3) there is a geometric phase associated with each $|\phi_{\lambda}^{ad}\rangle$, which is dependent on the path on the $\vec{R}$ space. The geometric phases are carried to $\underline{\underline{F}}$ and $\underline{\underline{G}}$; (4) in addition, there can be problematic adiabatic PES character which can be discontinuous and non-smooth. Overall, these disadvantages can render quantum chemistry calculations involving adiabatic electronic states to fail or produce incorrect results.

2.1.1 Diabatic States

To alleviate the disadvantages of adiabats, we can use another set of electronic states, diabats ($\{\phi^d\}$), to resolve $\hat{H}$. Diabats are related to adiabats by a unitary transformation, or more specifically, orthogonal transformation in this thesis:

$$|\phi_a^d\rangle = \sum_{\lambda} U_{a\lambda} |\phi_{\lambda}^{ad}\rangle \quad (2.16)$$

All the entities in this equality are functions of $\vec{R}$. Since the orthogonal transformation rotates the adiabats among themselves, the two sets of states span the same space. The identity operator of the electronic states constructed using the two sets of states is strictly equivalent:

$$\sum_a |\phi_a^d\rangle \langle \phi_a^d| = \sum_{\lambda} |\phi_{\lambda}^{ad}\rangle \langle \phi_{\lambda}^{ad}| \quad (2.17)$$

Therefore, we can resolve $\hat{H}$ using the diabats in a similar way as using the adiabats in Eq. 2.10:

$$\begin{aligned} \hat{H} &= \sum_{a,b} |\phi_a^d\rangle \langle \phi_a^d| \hat{H} |\phi_b^d\rangle \langle \phi_b^d| \\ &= \sum_{a,b} |\phi_a^d\rangle \langle \phi_a^d| \hat{T}_N + \hat{H}_{elec} |\phi_b^d\rangle \langle \phi_b^d| \\ &= \sum_{a,b} |\phi_a^d\rangle \left(\langle \phi_a^d| \hat{T}_N |\phi_b^d\rangle + \langle \phi_a^d| \hat{H}_{elec} |\phi_b^d\rangle \right) \langle \phi_b^d| \\ &= \sum_{a,b} |\phi_a^d\rangle \left(\langle \phi_a^d| \hat{T}_N |\phi_b^d\rangle + V_{ab}(\vec{R}) \right) \langle \phi_b^d| \end{aligned} \quad (2.18)$$

where the definition of $V_{ab} = \langle \phi_a^d| \hat{H}_{elec} |\phi_b^d\rangle$ is a function of $\vec{R}$. Since each diabate is a (linear combination) mixture of adiabats (eigenstates of $\hat{H}_{elec}$), diabats are not eigenstates of $\hat{H}_{elec}$

and hence $\underline{\underline{V}}$ is in general a non-diagonal real-valued symmetric matrix. Obviously,

$$\underline{\underline{V}}(\vec{R}) = \underline{\underline{U}}(\vec{R})^\dagger \underline{\underline{E^{ad}}}(\vec{R}) \underline{\underline{U}}(\vec{R}), \quad (2.19)$$

an orthogonal transformation of the diagonal $\underline{\underline{E^{ad}}}$ matrix. Since $\{E^{ad}(\vec{R})\}$ have the physical significance as adiabatic potential energy surfaces, their transformation, $\underline{\underline{V}}(\vec{R})$ is called the diabatic potential matrix.

But we have not defined diabats yet, since we have not defined the transformation matrix $\underline{\underline{U}}$. Diabats are defined as the electronic states that smoothly evolve with respect to nuclear configuration variation, i.e., they should smoothly parametrically depend on $\vec{R}$. The $\underline{\underline{U}}(\vec{R})$ matrix should generate the smooth $\{|\phi^d(\vec{R})\rangle\}$ from the potentially non-smooth $\{|\phi^{ad}(\vec{R})\rangle\}$. This itself means that the $\underline{\underline{U}}(\vec{R})$ must be non-smooth when some of $\{|\phi^{ad}(\vec{R})\rangle\}$ are non-smooth. The smoothness requirement on the diabats is to ensure

$$\langle \phi_a^d | \hat{T}_N | \phi_b^d \rangle \approx \delta_{ab} \hat{T}_N \quad (2.20)$$

i.e., the nuclear kinetic operator only acts on the nuclear wave packet, with negligible effect from the $\vec{R}$ -dependence of the diabats.

Following the same derivation in Eq. 2.10,

$$\langle \phi_a^d | \hat{T}_N | \phi_b^d \rangle = -\frac{1}{2} G_{ab}(\vec{R}) - i \vec{F}_{ab} \cdot \hat{\vec{P}} + \hat{T}_N \delta_{ab} \quad (2.21)$$

$\hat{\vec{F}}_{ab}$ has the explicit form of

$$\hat{\vec{F}}_{ab} = \sum_n \vec{n} \langle \phi_a^d | \frac{\partial}{\sqrt{M_n} \partial X_n} | \phi_b^d \rangle \quad (2.22)$$

where n indexes the nuclear degree of freedom, whose coordinates and unit vectors are labeled as X_n and $\vec{n}$, respectively. The mass-weighting is explicitly written. As long as the diabats

smoothly vary with respect to $\vec{R}$, $\frac{\partial}{\partial X_n} |\phi_b^d\rangle$ is finite, and $\frac{\partial}{\sqrt{M_n}\partial X_n} |\phi_b^d\rangle$ is a small quantity since M_n is at least 2000 in atomic units. Consequently, $\langle \phi_a^d | \frac{\partial}{\sqrt{M_n}\partial X_n} |\phi_b^d\rangle$ is a small quantity, i.e.,

$$\hat{\vec{F}}_{ab} \approx \vec{0} \quad (2.23)$$

The relation between $\underline{\underline{G}}$ and $\underline{\underline{R}}$ in Eq. 2.13 is derived without considering diabats or diabats, and therefore, this relation holds for diabats. This relation and $\hat{\vec{F}}_{ab} \approx \vec{0}$ means that $G_{ab} \approx 0$. We shall not forget that another fold of mass-weighting is hidden in $\hat{\vec{P}}$. Given all these considerations, Eq. 2.21 is reduced to the desired Eq. 2.20.

Clearly, the motivation of introducing diabats is to remove the trouble-making non-adiabatic coupling matrix $\underline{\underline{F}}$ from the resolved $\hat{H}$, and hence from the Schrödinger equation, time-dependent or time-independent. Strict diabats are those with $\frac{\partial}{\partial X_n} |\phi_b^d\rangle = 0$. This is a strong requirement that can only be satisfied when the space of electronic states is fully truly complete, i.e., when literally all electronic states (ground, excited, bound, scattering) are included in $\sum_\lambda |\phi_\lambda^{ad}\rangle \langle \phi_\lambda^{ad}|$ and $\sum_a |\phi_a^d\rangle \langle \phi_a^d|$. It is impractical to use such an ∞ -dimensional space in solving actual problems. Therefore, the definition is relaxed to $\frac{\partial}{\partial X_n} |\phi_b^d\rangle$ being not large. The diabats hence do not need to be strictly $\vec{R}$ -invariant; they only need to be $\vec{R}$ -smoothly-variant. The strict diabats are now relaxed to quasi-diabats [6, 35]. The negligence of $\vec{F}_{ab}$ and G_{ab} for the quasi-diabats does not induce error more significant than BOA for well separated diabats. Henceforth, all “diabat” means quasi-diabat.

Completeness matters, despite the infeasibility of achieving it fully. Hence, ideally, a finite set of electronic states is adequate to describe the problem at hand, though it is not truly complete. This is to guarantee that a set of diabats can be transformed to a set of diabats that preserve their characters (e.g., local excitonic states, a specific valence bond structure, etc.) in the range of $\vec{R}$ that is of interest. Although this means that a set of diabats change their character (suddenly or gradually), the changes are exchanges among the set, without exchanges with states out of the set. In other words, this set of diabats

are energetically well separated from the other adiabats. This is called the group Born-Oppenheimer approximation (GBOA), since the BOA is now applied to a group of states. Since the characters of the set of states are preserved, we can always find the diabats, each of which features certain characters. Therefore, the validity of GBOA is the foundation for finding diabats for adiabats. This process, we formally call diabaticization.

Now, if we look back at the diabats-resolved Hamiltonian in Eq. 2.8, it becomes

$$\hat{H} = \hat{T}_N + \sum_{a,b} |\phi_a^d\rangle V_{ab}(\vec{R}) \langle \phi_b^d| \quad (2.24)$$

$\hat{T}_N$ acts solely on nuclear wave packet, while the V_{ab} elements are functions of *vibrational* coordinates and couple different *electronic* states; the $\underline{V}$ matrix delivers the vibronic interaction. It is hence also called the vibronic matrix. From the adiabatic to the diabatic (ADT) representation, the vibronic coupling has been transferred from the $\hat{T}_N$ kinetic operator to the $\underline{V}$ potential operator.

In the diabatic representation, the whole vibronic state reads

$$|\Psi\rangle = \sum_a |\phi_a^d\rangle |\chi_a^d\rangle \quad (2.25)$$

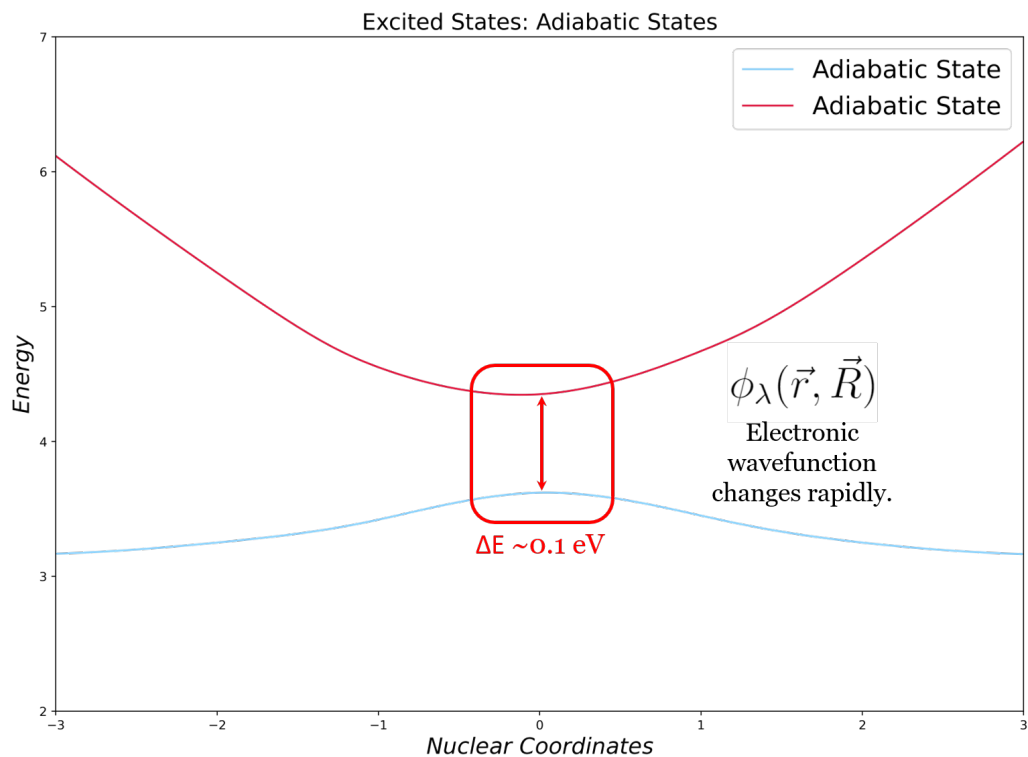
Now, $|\chi_a^d\rangle$ is the vibrational wave packet associated to the diabat $|\phi_a^d\rangle$. The operation of the Hamiltonian in Eq. 2.24 onto this vibronic state, we have

$$\hat{H} |\Psi\rangle = \sum_a |\phi_a^d\rangle \left(\hat{T}_N |\chi_a^d\rangle + \sum_b V_{ab}(\vec{R}) |\chi_b^d\rangle \right), \quad (2.26)$$

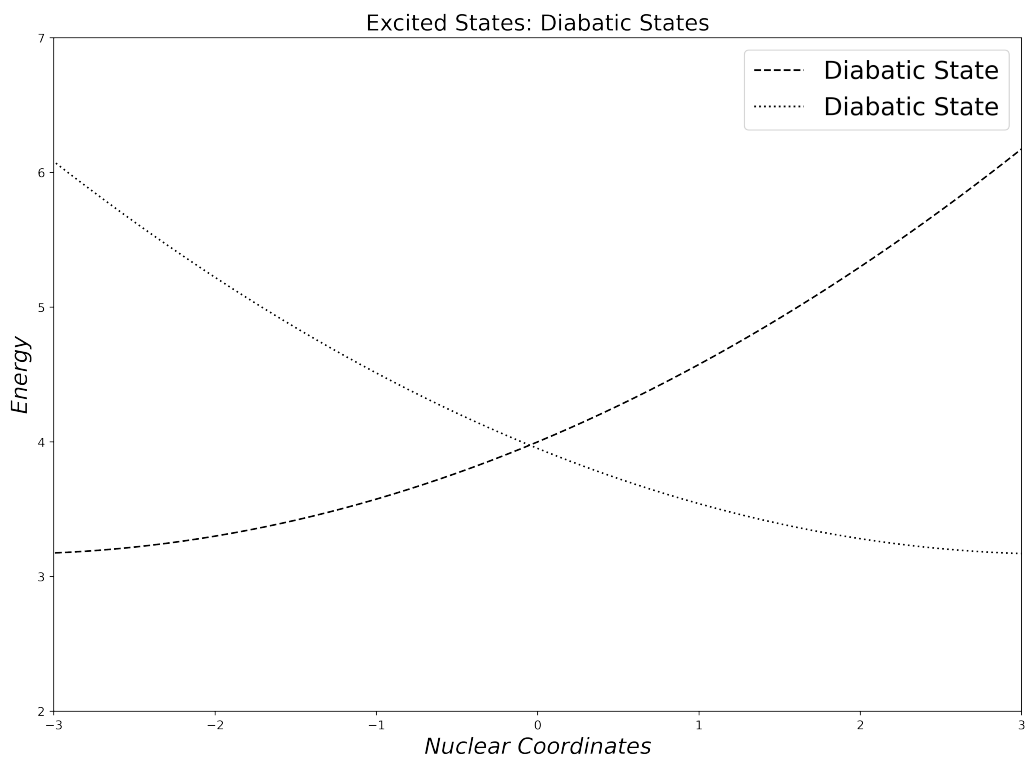
with $\hat{T}_N$ and $V_{aa}(\vec{R})$ determining the evolution of nuclear wave packet on the diabatic PES of $|\phi_a^d\rangle$, while $V_{a \neq b}(\vec{R})$ and $V_{aa}(\vec{R}) - V_{bb}(\vec{R})$ controlling the transition of the wave packets between different diabatic PESs.

The diabatic representation features the following advantages. First, removal of the singular, not-easy-to-calculate NACME and removal of the geometric phase problem [3]. The electronic states (diabats) are associated with clear character (charge-transfer, local excitation, biradical, covalent, essentially all terms that you can think of based on valence bond theory), which facilitates discussion and understanding. This character preservation makes the diabatic basis highly useful in simulating non-adiabatic chemical processes [3]. Diabats are $\vec{R}$ -differentiable everywhere and maintain smooth character throughout, and so are $V_{ab}(\vec{R})$. This diabatic potential matrix, a globally smooth function of nuclear configuration, determines the force that the vibrational wave packet receives and the transitions of the wave packet between different states [35]. The smoothness also allows us to describe $V_{ab}(\vec{R})$ as polynomial expansions in $\vec{R}$.

The diabatic representation features the following disadvantages. Diabats lose the meaning of electronic energy levels, as they are not eigenstates of $\hat{H}_{elec}$. Transitions between diabats can occur in a larger space of $\vec{R}$ than between adiabats. It takes post-processing (diabatization) to obtain diabats from adiabats [10]. To obtain $V_{ab}(\vec{R})$ as a polynomial expansion, we need diabats on a grid of $\vec{R}$. Clearly, the pros and cons of diabats are the cons and pros of adiabats. The two representations, as compared in Figure 2.3, complement each other. While adiabats might be more intuitive to calculate, the mathematical advantageousness of diabats lends itself to be more appropriate for addressing non-adiabatic chemistry problems. Therefore, diabats are worthwhile to calculate despite their own shortcomings.



(a) Adiabatic states, complicated to represent.



(b) Diabatic states, comparatively easier to represent.

Figure 2.2: Comparison of adiabatic and diabatic states for a system with two electronic excited states.

2.1.2 Our Diabatization Scheme

Our diabatization scheme follows the philosophy that the diabats at two nuclear configurations that are finite-differentiably close have a maximum overlap. The detailed scheme is as follows.

- i. We first choose a nuclear configuration as the reference geometry. If we are handling a photoelectron ionization problem, then the optimized structure of the neutral molecule is a good option.
- ii. Several adiabats at the reference geometry relevant to the problem at hand are obtained.
- iii. Several vibrational modes relevant to the problem are selected. They can be normal modes or internal motions, such as bond bonding or bond stretching.
- iv. Along each of the several modes, a mesh grid of displacements is constructed. We then distort the system away from the reference geometry on the grid. Adiabats are obtained on each grid point.
- v. Let's say for two neighboring grids $\vec{q}_0$ and $\vec{q}_1$, we have obtained the diabats at $\vec{q}_0$, which are labeled as $\{|\phi^{d,0}\rangle\}$, but still need to obtain the diabats at $\vec{q}_1$, labeled as $\{|\phi^{d,1}\rangle\}$, from the adiabats at $\vec{q}_1$, labeled as $\{|\phi^{a,1}\rangle\}$.
- vi. For each $|\phi_i^{d,0}\rangle$, we construct a projection operator: $\hat{P}_i = |\phi_i^{d,0}\rangle\langle\phi_i^{d,0}|$, calculate its matrix elements for the adiabats $\{|\phi^{a,1}\rangle\}$: $P_{\lambda\mu}^i = \langle\phi_\lambda^{a,1}|\hat{P}_i|\phi_\mu^{a,1}\rangle$, diagonalize $\underline{\underline{P}}^i$ to obtain the eigenvector with the largest eigenvalue. This eigenvector gives the linear combination of $\{|\phi^{a,1}\rangle\}$ that has the maximum overlap with $|\phi_i^{d,0}\rangle$. We call it preliminary diabats $|\tilde{\phi}_i^{d,1}\rangle$ at $\vec{q}_1$.
- vii. A set of preliminary diabats $\{|\tilde{\phi}^{d,1}\rangle\}$ are then obtained. Since they are obtained as eigenstates of different projection operators, $\hat{P}_1$, $\hat{P}_2$, etc., they are not orthogonal to

each other. We then symmetrically orthonormalize the preliminary diabats to have the final set of true diabats at $\vec{q}_1$, $\{|\phi^{d,1}\rangle\}$. The phases of those diabats have been set to give $\langle \phi_i^{d,0} | \phi_i^{d,1} \rangle > 0$.

- viii. Along a path on the grid starting from the reference geometry, for the first displacement from the reference geometry, the adiabats at the reference geometry are used as $\{|\phi^{d,0}\rangle\}$. For all the following grids, the diabats of the previous grid are $\{|\phi^{d,0}\rangle\}$.

Finally, once the set of true diabats $\{|\phi^{d,i}\rangle\}$, are found for all displacements on the grid of $\vec{R}$, vibronic coupling constants for each normal mode can at this moment be determined by Taylor series expansion [43, 47]:

$$V_{ab}(\vec{q}) = V_{ab}(\vec{q}_0) + \sum_i^M V_{ab}^i \vec{q}_i + \frac{1}{2} \sum_{i>j}^M V_{ab}^{ij} \vec{q}_i \vec{q}_j + \dots \quad (2.27)$$

where the first-order Taylor coefficient can for example be evaluated using finite difference numerical differentiation:

$$V_{ab}^i = \frac{V_{ab}(\vec{R}_{\vec{q}_0} + d\vec{q}_i) - V_{ab}(\vec{R}_{\vec{q}_0} - d\vec{q}_i)}{2d\vec{q}_i} \quad (2.28)$$

2.2 Building Wavefunctions

2.2.1 Multi-configuration Self Consistent Field (MCSCF)

To improve on the traditional single Slater determinant (SD) Hartree-Fock (SCF) method, when a linear combination of SDs are used to construct the reference wave function instead, this is MCSCF. MCSCF will proceed to optimize both the configuration state function (CSFs) coefficients and molecular orbital basis simultaneously [48]. Comparatively, CI only touches upon optimizing the coefficients, and leaves the orbitals from SCF untouched. This means MCSCF wavefunctions are quite balanced and more flexible than HF or CI, but of course more costly. Only a small number of SDs arising from flexible occupations of a small number (e.g., 16) of electrons in a small number (e.g., 16) of orbitals can be included in a MCSCF treatment. The set of electrons (e) and orbitals (o) form the active space, which is usually labeled as, e.g., 16o16e, 8o10e, etc. The few SDs that contribute most significantly to the electronic states of interest shall be captured in the MCSCF treatment, i.e., non-dynamic correlation should have been treated by MCSCF [49]. The correlation from the leftover large number of SDs can be recovered by their inclusion in a PT manner, i.e., dynamic correlation should have been treated by PT. MCSCF + PT results in the generic MCPT composite treatment. The relation between MCPT2 and MCSCF is like the relation between MP2 and HF.

2.2.2 CASSCF, RASSCF, CASPT, and GMC-QDPT

A *mon*e active space can be further partitioned into several (say N) subspaces: $m_1o_1n_1e$, $m_2o_2n_2e, \dots, m_No_Nn_Ne$. If we do not do this further partition, then MCSCF becomes complete active space SCF (CASSCF) [45]. If we allow several subspaces, with restrictions on the electron promotions between the subspaces, then MCSCF becomes restricted active space SCF (RASSCF) [50]. A CASSCF wave function is invariant with respect to a free rotation

(i.e., mixture) among all the active orbitals, while a RASSCF wave function is invariant with respect to the rotations among the respective subspaces. All calculations reported in this thesis are based on CASSCF calculations¹. In principle, the method of this thesis can be generalized to RASSCF calculations. The PT treatment based on CASSCF calculation is called CASPT calculation [52]. Without further specification, all PT treatments in this thesis are PT2 treatments. There are many ways to add the PT treatment. The specific PT treatment we use is called the general multi-configurational quasi-degenerate perturbation theory (GMC-QDPT), developed by Nakano and coworkers [31, 32, 33].

Diabatic Molecular Orbitals

In our diabaticization scheme, we first need to express the diabats using a set of MOs that are largely invariant with respect to $\vec{R}$. For obvious reason, this set of MOs are called diabatic MOs (DMOs) [35, 53]. DMOs are orthogonal transformation of canonical MOs (CMOs). CMOs can be viewed as adiabatic MOs, as they are eigenorbitals of the Fock operator at individual $\vec{R}$. The procedure to obtain DMOs from CMOs resembles the procedure to obtain diabats from adiabats in Section 2.1.2. Since CASSCF adiabats, and consequently GMC-QDPT adiabats, are invariant with respect to this orbital rotation among the active space, the GMC-QDPT adiabats can be represented using DMOs, i.e., each adiabat is a linear combination of SDs in DMOs. This set of SDs are labeled as DMO-SDs.

Now, $\{\phi^{d,0}\}$ and $\{\phi^{a,1}\}$ in Step vi. in the diabaticization procedure are expressed using the DMO-SDs. The $\langle \phi_i^{d,0} | \phi_\lambda^{a,1} \rangle$ overlap is then expressed as a dot product of the linear coefficients of the diabat and adiabat in the DMO-SDs and are easy to obtain. This diabat-adiabat overlap matrix gives the $\underline{\underline{P}}^i$ matrix. Then the ingredients of the diabaticization recipe are all in place. Diabats are then obtained.

¹Within our template GAMESS calculations, we actually invoke `ORMASSCF` (Occupation Restricted Multiple Active Space) calculations [51], but since it is set to a single active space, CASSCF wavefunctions are produced.

2.2.3 Active Space Design

Active space classifies the explicit region of orbitals and electrons that will be explored in the quantum chemistry calculation² [48, 50]. It can be sectioned into active orbitals and electrons, and an inactive subset. The active electrons populate the active orbitals, however the active orbitals can include unoccupied orbitals, where "unoccupied" refers to those that are unoccupied in the single determinant SCF, but included in the active space for CAS. If they are part of the inactive subset³, then they are known as frozen (or core) orbitals if they are doubly occupied. Virtual orbitals are those that are unoccupied in all CSFs.

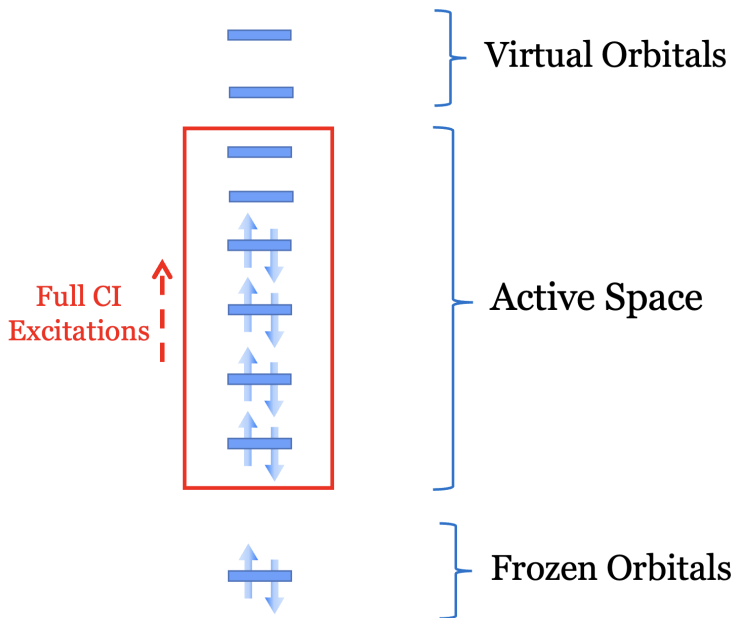


Figure 2.3: Quantized energy level diagram showing orbitals and electrons in a standard water molecule 6o8e full valence active space. 6 orbitals and 8 electrons are included. The levels are tripartitioned into frozen, active, and virtual orbitals.

Designing the active space is a challenging problem [54]. In order to nail down the active space, chemical intuition and strategic assessment of orbitals is needed. Full valence active space is the most ideal situation. As the dimensions of the active space increases, calculations frequently fail to converge, so adjustments and restrictions are needed.

²We will only discuss the active space as it applies to the CASSCF case in this thesis.

³PT treatments can have excitations from inactive and into virtual orbitals.

2.3 Spin Orbit Coupling, Jahn-Teller Effect, and Pseudo-Jahn-Teller Effects

2.3.1 Spin Orbit Coupling

There have been a good number of review articles and monographs on spin-orbit coupling (SOC) [55, 56]. We will discuss SOC here in detail, but only give a brief conceptual introduction. In an atom, each electron moves around the nucleus, and possesses a spin magnetic dipole moment. In the reference frame of the electron, it is the nucleus that moves around. This motion of the positively charged nucleus creates a magnetic field onto the electron. The interaction between the magnetic field and the electron's spin magnetic dipole gives a term in the $\hat{H}_{elec}$. Only the angular motion of the nucleus around the electron creates the magnetic field. Therefore, the aforementioned interaction involves a dot product between the nuclear spatial angular momentum and the electron's spin angular momentum.

Coming back to the conventional reference frame of the nucleus, the nuclear spatial angular momentum around the electron becomes the electron's spatial angular momentum around the nucleus. The dot product is between the electron's spatial angular momentum (*orbit*) around the nucleus and the electron's *spin* angular momentum. Therefore, this term is called spin-orbit coupling, which in general has a form of

$$\hat{H}_{SO} = \zeta(Z, r) \hat{\vec{l}} \cdot \hat{\vec{s}} \quad (2.29)$$

with $\hat{\vec{l}}$ being the orbital angular momentum operator, $\hat{\vec{s}}$ the spin angular momentum operator, and $\zeta(Z, r)$ a monotonically decreasing function of the nucleus-electron distance r and monotonically increasing function of the nuclear charge Z . $\zeta(Z, r)$ determines that the spin-orbit coupling is enhanced by a shorter nucleus-electron distance (r) and/or a higher nuclear charge (Z). For a polyatomic multi-electron system, we simply sum up SOC between

all electron-nucleus pairs

$$\hat{H}_{SO} = \sum_{A,i} \zeta(Z_A, r_{iA}) \hat{\vec{l}}_{iA} \cdot \hat{\vec{s}}_i \quad (2.30)$$

where i and A index electrons and nuclei, respectively. The rest of the equation shall be self-explained.

Electron-electron SOC exists and can be viewed as a screening of the nucleus-electron SOC, just like electron-electron electrostatic repulsion screens off nucleus-electron attraction. Therefore, we do not consider electron-electron SOC explicitly, but conceptually noting that it reduces the magnitude of $\zeta(Z, r)$.

In the conventional position representation, $\hat{\vec{l}}$ contains the imaginary unit i , which is inherited by $\hat{H}_{SO}$. Therefore, matrix elements of real-valued electronic states, e.g., $\langle \phi_a^d | \hat{H}_{SO} | \phi_b^d \rangle$ or $\langle \phi_\lambda^a | \hat{H}_{SO} | \phi_\mu^a \rangle$ are generically complex-valued. The dependence of $\hat{H}_{SO}$ on nuclear positions through r_{iA} and $\hat{\vec{l}}_{iA}$ determines that the matrix element

$$V_{ab}^{SO}(\vec{R}) = \langle \phi_a^d | \hat{H}_{SO} | \phi_b^d \rangle \quad (2.31)$$

is a function of $\vec{R}$. Just like $V_{ab}(\vec{R})$ of the electrostatic Hamiltonian, $V_{ab}^{SO}(\vec{R})$ is a smooth function and can be expressed as a polynomial in $\vec{R}$. The $\underline{\underline{V}}^{SO}$ matrix delivers the spin-orbit vibronic coupling, as it arises from *spin-orbit* interaction, is a function of *vibrational* coordinates, and couples different *electronic* states.

The Z -dependence of $\hat{H}_{SO}$ determines that SOC is significant for heavy elements. Specifically, SOC scales with Z^2 [55]. Correspondingly, spin-orbit vibronic interaction is substantial for heavy element compounds [57]. It is unignorable. While this is the overall impression, SOC is also of importance for light-element chemistry, whenever inter-system crossing occurs. Another concern is degeneracy between electronic states. Due to the hermiticity and purely imaginary nature of $\hat{\vec{l}}$, diagonal V_{aa}^{SO} or $V_{\lambda\lambda}^{SO}$ elements are strictly zero for real-valued

diabats or adiabats. Therefore, SOC has its maximum impact between degenerate or pseudo-degenerate electronic states. And the impact is diminished upon the increase of the energy gap between two states.

D_{3h}	E	$2C_3$	$3C'_2$	σ_h	$2S_3$	$3\sigma_v$
A'_1	1	1	1	1	1	1
A'_2	1	1	-1	1	1	-1
A''_1	1	1	1	-1	-1	-1
A''_2	1	1	-1	-1	-1	1
E'	2	-1	0	2	-1	0
E''	2	-1	0	-2	1	0

Table 2.1: Character Table and IRREPs for the D_{3h} Point Group. Since the first order SOC requires a spatially degenerate term symbol with non-zero spin, we focused on ^2E -type term symbols in our study [39, 58].

2.3.2 Jahn-Teller

Any nuclear configuration with $C_{n>2v}$, $D_{n>2}$, $D_{n>2d}$, $D_{n>2h}$, $S_{2n>2}$, T , T_h , T_d , O , O_h , I , I_h symmetries can give degenerate electronic states [57, 59]. The point groups of the symmetries are non-Abelian groups and therefore, they have multi-dimensional irreducible representations (IRREPs). An example of IRREPs for the D_{3h} point group is shown in Table 2.1. The electronic Hamiltonian of a molecule has the same symmetry as the nuclear framework, and the eigenstates of the Hamiltonian must transform as the IRREPs of the corresponding point group. Those that transform as components of the same multi-dimensional IRREP must be degenerate. This argument applies to canonical molecular orbitals, which are eigenstates of the Fock operator, and multi-electronic eigenstates of the electronic Hamiltonian. The e_g - t_{2g} two-over-three MO diagram of O_h transition metal compounds is an example of the former. Putting, say, two electrons on the t_{2g} shell, the t_{2g}^2 occupation scheme gives $^3T_{1g}$, $^1T_{2g}$, ^1E_g , and $^1A_{1g}$ term symbols, which are examples of the latter. Those degenerate states are called spatially degenerate states, as the degeneracy arises from the spatial symmetry of the system.

The spatially degenerate electronic states shares one feature: the degenerate MOs are not equally populated. This is easy to understand. There is only one way to equally populate all degenerate MOs, and this only option gives a non-degenerate state. For instance, the ${}^1A_{1g}$ term symbol consists of the following linear combination of Slater determinants:

$$|{}^1A_{1g}\rangle = \frac{1}{\sqrt{3}} (|t_{2g,xy}\bar{t}_{2g,xy}\rangle + |t_{2g,yz}\bar{t}_{2g,yz}\rangle + |t_{2g,xz}\bar{t}_{2g,xz}\rangle) \quad (2.32)$$

An overhead bar is used to label spin orbitals with β spinor, while those without the bar have α spinor. Clearly, on average, each of the three t_{2g} orbitals is equally occupied by $\frac{2}{3}$ electrons in this state. Electron(s) in each of the three orbitals exerts a distorting force along each of the three orthogonal tetragonal axes of the octahedron. The equal $\frac{2}{3}$ average population means the equal magnitude of the three distorting forces. The tetragonal distortions cancel each other, and the framework remains its O_h symmetry, with some expansion/contraction only. I.e., the electrostatic force of the electronic state with equal populations on the degenerate orbitals is of the same symmetry as the nuclear framework. This picture can be straightforwardly generalized to the other symmetries.

Taking a degenerate term symbol, ${}^3T_{1g}$, for example, one of the 9 component states takes the following one-SD form:

$$|{}^3T_{1g,R_y,M_S=1}\rangle = |t_{2g,xy}t_{2g,yz}\rangle \quad (2.33)$$

with one electron in each of the $t_{2g,xy}$ and $t_{2g,yz}$ spatial orbitals, leaving the $t_{2g,xz}$ orbital empty. The balanced electrostatic force in the ${}^1A_{1g}$ state in Eq. 2.32 becomes unbalanced for the ${}^3T_{1g}$ states in Eq. 2.33. Now, the force is of D_{4h} symmetry, a subgroup of O_h , and the y -axis is the 4-fold axis. This force certainly distorts the structure and lowers its symmetry from O_h to D_{4h} . The symmetry-lowering can be a compression or an elongation along one of the three tetragonal axes (y -axis in this case) [57].

The arguments above can be generalized to the other high symmetries, as the origin

of the symmetry-lowering is the same, that spatially degenerate orbitals are not equally populated. Such a symmetry-lowering, brought by symmetry-induced degeneracy, is called the Jahn-Teller (JT) distortion. While it may appear paradoxical that symmetry induces a force to distort that very symmetry, this phenomenon highlights the beauty and complexity of nature. It is important to note that the Jahn-Teller distortion arises from electrostatic interaction rather than covalent interactions. It is a symmetry-induced vibronic interaction and features the conical intersection topology of adiabatic PESs.

2.3.3 Pseudo-Jahn-Teller

Similar symmetry-lowering distortion can also occur for non-degenerate electronic states. One typical example is NH_3 , which, in its ground state, distorts from the trigonal planar D_{3h} structure to the trigonal pyramidal C_{3v} structure. The ground electronic state of NH_3 in its C_{3v} form is dominated by a Slater determinant, with the occupation scheme of $[\text{core}](1a_1)^2(e)^4(2a_1)^2$, with the N-H bonding character in the $1a_1$ orbital and the two e_1 component orbitals. The $2a_1$ is the lone-pair orbital. In its D_{3h} transition state structure between the two mirror images of the C_{3v} equilibrium structures, the ground state Slater determinant has the occupation scheme of $[\text{core}](1a'_1)^2(e')^4(2a'_1)^2$.

The two sets of valence MOs at the distorted and undistorted structures are compared in Figure 2.4. In the figure, we see very nice correlation between the orbitals along the pyramidalization. The most noticeable change is the energy lowering of the lone-pair orbital (a''_2 in D_{3h} and $2a_1$ in C_{3v}). This orbital energy lowering leads to the overall energy lowering of the ground state and hence drives the pyramidalization. This energy lowering of the lone pair orbital comes from its mixing with the $2a'_1$ antibonding orbital at the D_{3h} structure. This mixing is forbidden at the D_{3h} symmetry since the two orbitals are of different IRREPs (a''_2 vs a'_1). However, the mixing is turned on for the symmetry lowering in the pyramidalization: in C_{3v} , they are both a_1 orbitals. The bonding mixing pushes down the lone pair energy,

meanwhile pushing up the antibonding unoccupied $3a_1$ orbital.

This mixing (coupling) between electronic states (exemplified as one-electron orbitals here) is induced by a symmetry-lowering vibration. It is a type of vibronic coupling. This type of vibronic interaction shares the same symmetry-lowering feature with the Jahn-Teller interaction, but it involves non-degenerate electronic states. Therefore, it is called the pseudo-Jahn-Teller interaction. All effects induced by this interaction are called pseudo-Jahn-Teller (pJT) effects [59]. Despite the resemblance to the Jahn-Teller interaction, pseudo-Jahn-Teller interaction has a different origin. The symmetry-lowering-induced orbital mixing is intrinsically a covalent interaction, versus the electrostatic origin of the Jahn-Teller interaction. Covalent interaction arises from orbital overlaps.

When two states are closer in energy at a high symmetry configuration, the direct product of their IRREPs contains the IRREP of a symmetry-lowering mode, and the distortion along this mode results in substantial overlap between the two states, then the two states have a strong pseudo-Jahn-Teller interaction along this mode. The three conditions above parallel the three rules of covalent interaction between orbitals of two fragments: (1) energy close; (2) symmetry matching; (3) maximum overlap. These three conditions are nothing more than the first-order wave function correction and second-order energy correction of the Rayleigh-Schrödinger Perturbation Theory [60].

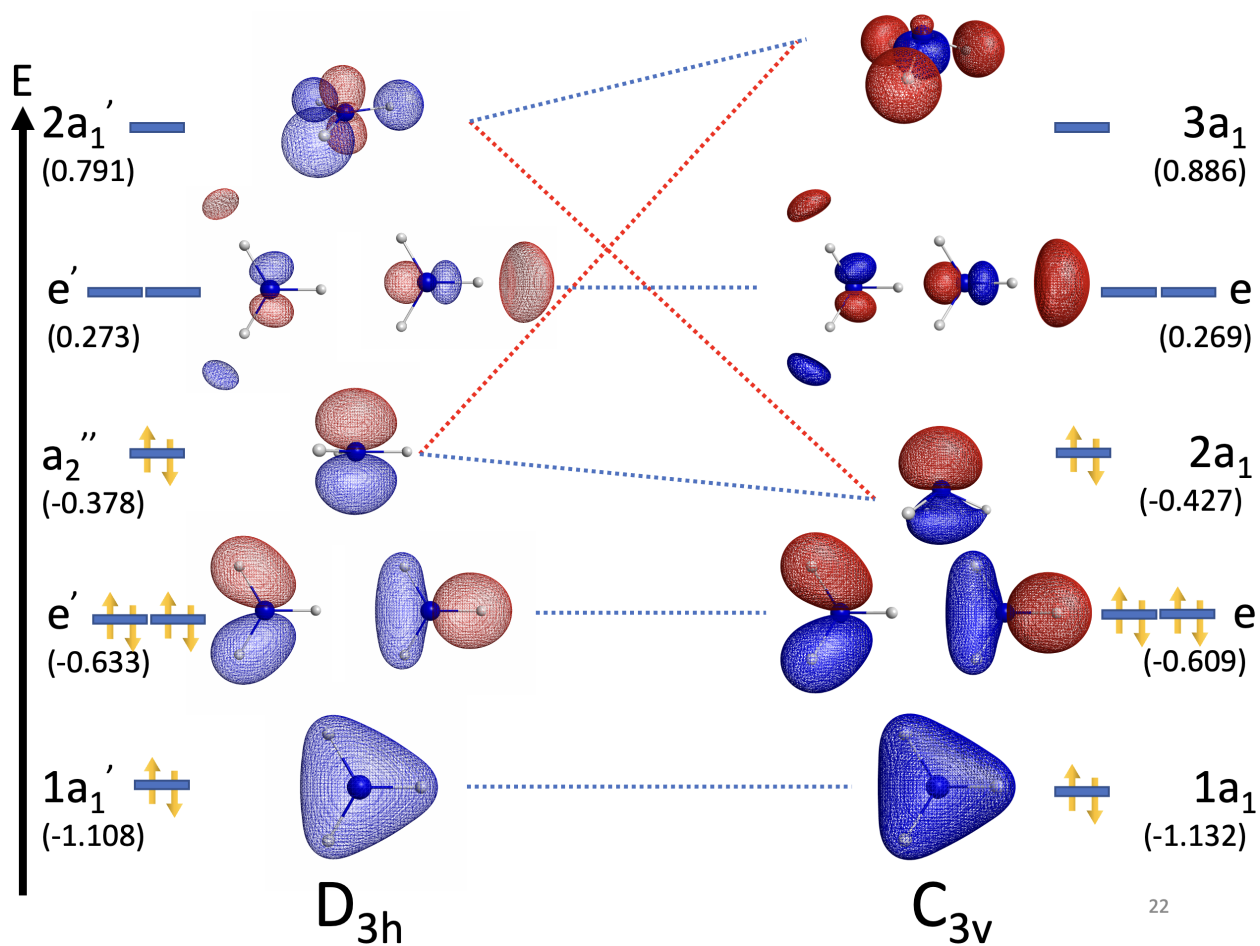


Figure 2.4: Comparison of D_{3h} and C_{3v} NH_3 orbitals, the HOMO is a_2'' and undergoes Jahn-Teller distortion to stabilize as C_{3v} 's $2a_1$ lowers in energy. The C_{3v} NH_3 has bonding orbitals comparatively lower in energy than D_{3h} . Mixing interactions are highlighted in red.

2.3.4 The Interplay between Spin-Orbit and (Pseudo-)Jahn-Teller

As mentioned above, SOC is the strongest when the SO coupled states are degenerate. The orbital rotation of an electron corresponds to an electron hopping (or “excitation”) from one real-valued orbital to another real-valued orbital that have different orientations, e.g., from a p_x orbital to a p_y orbital, from a π orbital to a σ orbital, etc. If this hopping (or “excitation”) does not cost energy, then this orbital rotation is easier to occur. That is why SOC is stronger when the two states coupled by the spin-orbit interaction are closer in energy, and the strongest when they are degenerate. An electronic state with all spatial orbitals being real-valued cannot have SOC in itself. This is due to the two characters of the orbital angular momentum operator: hermiticity and being purely imaginary in position representation. That is why SOC only occurs between different states of real-valued spatial orbitals. The two states may have similar energies, that is why we put quotation marks on “excitation” above. When an orbital angular momentum operator acts on a real-valued spatial orbital, if the orbital points in the same direction as the operator, e.g., $\hat{l}_z |p_z\rangle$, both are along the z direction, the operation is equal to annihilation. Consequently, there is no SOC. Only when the orbital and the operator are perpendicular, e.g., $\hat{l}_x |p_z\rangle$, the orbital is rotated about the axis of the operator. That is the situation when the SOC matrix element becomes large. This rule is called the El-Sayed’s rule [61]. When the two electronic states corresponding to this hopping between orbitals are closer in energy, this off-diagonal coupling is stronger. The degeneracy of electronic states hence bridges SOC and JT, and the pseudo-degeneracy bridges SOC and pJT.

SOC and JT can mutually quench each other [2, 40]. For instance, the SOC-induced hopping between orbitals results in mixing of degenerate states that are differently oriented, e.g., the xz, yz, xy spatial components of the $|^3T_{1g}\rangle$ term symbol mentioned above. If it is an equal mixture, then the JT electrostatic forces along different directions are averaged to zero, resulting in no JT distortion. Vice versa, along the JT distortion, the degeneracy between

the spatial components is lifted, which weakens their SOC. The same argument applies to the relation between SOC and pJT, as the gap enlargement by pJT distortion also quenches SOC; SOC at the undistorted structure enlarges the gap between the pJT-coupled states, and then quenches pJT.

SOC and JT can also mutually enhance each other. The JT distortion splits originally degenerate states, but may bring those split states closer to some other states, and hence enhance SOC between the two sets of states. The energy lowering induced by the stronger SOC retrospectively furthers the JT distortion. The same argument applies to the possible mutual enhancement between SOC and pJT. Overall, SOC is intimately connected to JT and pJT. This is especially true for high symmetry heavy element compounds, where the three effects are all potentially pronounced. For those compounds, SOC and vibronic coupling (JT + pJT) shall be treated on the same footing.

2.4 Spectroscopic Simulations

Discussion of electronic spectroscopy in quantum chemistry inevitably relates to excited states and the traditional Franck-Condon principle, which provides an understanding for the mechanism of electronic transitions and how they lead to spectral lines, as conceptualized in Figure 2.5. The assumption under the Franck-Condon approximation is that the photon-electron interaction occurs more rapidly than electronic and nuclear motions [13, 15, 44]. This assumption is reasonable for linear optics. The direct overlap of the ground state with the excited state produces an intensity in absorption and emission spectra [62]. This method was not directly employed in this thesis, but serves as a good foray into the traditional quantum dynamical methods in simulating spectroscopy. Typically, FC methods are coupled to efficient TD-DFT schemes, such as TD-DFT/VGFC (Vertical Gradient Franck-Condon) practiced in ORCA [16, 44, 63].

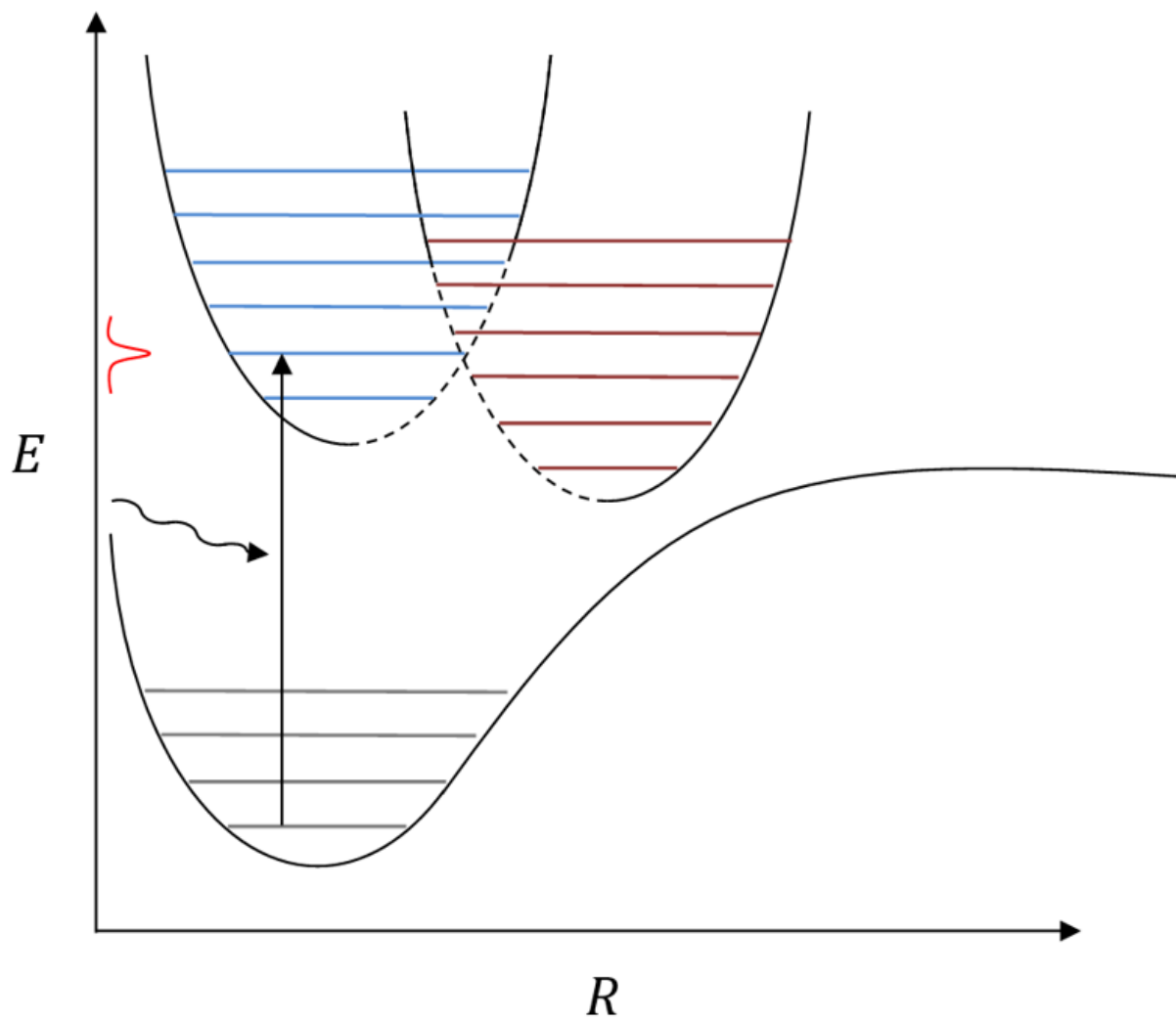


Figure 2.5: Common demonstration of electronic transition and band-shape intensities in spectroscopic profiles using the Franck-Condon Principle.

2.4.1 MCTDH

MCTDH is specifically a numerical approximation algorithm for solving the Time Dependent Schrödinger Equation (TDSE) for multi-component systems in many dimensions efficiently [64, 65]. It satisfactorily describes evolution of nuclear wave packets on several PESs. Importantly, MCTDH can propagate wavepackets and generate time-autocorrelation functions for very accurate vibrationally-resolved spectroscopy simulations [66]. With these strengths, MCTDH is the benchmark method to be compared with the novel VECC method. However, MCTDH has a high learning curve and is computationally inconvenient to non-specialists. It can be tricky to configure the right parameters such as the degrees of freedom, single-particle functions (SPFs) and basis functions (BFs) used to represent the wavefunction. In this project, the Heidelberg MCTDH86 package was used, and consultation with Prof. Dieter Meyer of MCTDH was needed throughout this project to resolve many complications.

2.4.2 VECC

VECC is a newly published Nooijen lab approach to non-adiabatic dynamics simulations based on a vibronic model framework by Bao, Nooijen, and Raymond [8]. Similar to MCTDH, it can be used to extract the expected suite of time-dependent properties for spectroscopy, with the added bonus of statistical mechanics. Although VECC is less systematic compared to MCTDH, VECC’s user-friendliness is superior, since it requires vastly fewer user input parameters (limited to simple on/off switches) and it is also advantageous in its (in theory) polynomial $O(n^2)$ time scaling, where n is the number of normal modes degrees of freedom (d.o.f.). These characteristics are summarized in Table 2.2.

VECC uses a second quantization mixed coupled cluster (CC) and configuration interaction (CI) ansatz to parameterize $|\psi(\tau)\rangle$:

$$|\psi(\tau)\rangle_b = \sum_x e^{\hat{T}(\tau)} \hat{Z}_x(\hat{\tau}) |x, 0\rangle \quad (2.34)$$

In 2.34, cluster operator $\hat{T}(\tau)$, and CI operator $\hat{Z}_x$ can be used to set the truncation order (of the infinite expansion), which determines the accuracy of the coupled cluster calculation. Currently in VECC, only up to $T = 1$ (single excitations) and $Z = 3$ (CISDT) is supported. Equation 2.35 is an example where $T = 1$ and $Z = 2$.

$$\hat{T} = \sum_i t^i(\tau) \hat{a}_i^\dagger ; \quad \hat{Z}_x = z_x^0 + \sum_i z_x^i \hat{a}_i^\dagger + \frac{1}{2} \sum_{ij} z_x^{ij} \hat{a}_i^\dagger \hat{a}_j^\dagger + \dots \quad (2.35)$$

Approach	Input	Computational Scaling	Thermal	User Friendly
MCTDH	Vibronic Hamiltonian	Exponential	No	No
VECC	Vibronic Hamiltonian	$O(n^2)$	Yes	Yes

Table 2.2: Comparison of the MCTDH and VECC approaches, highlighting their respective advantages and capabilities. Both methods can use the same vibronic Hamiltonian input. The polynomial $O(n^2)$ scaling is in terms of n normal modes. VECC can calculate thermal properties and is configured to only use simple on/off flags.

2.5 Computational Hardware and Software

2.5.1 GAMESS

General Atomic and Molecular Electronic Structure System (GAMESS) is the quantum chemistry package of choice in this project [67]. Specifically, the GAMESS-US (American) version maintained by the Gordon lab with *gmcpt* module by Zeng is the main engine for running the CASSCF/MRPT (SOC) calculations. The general pipeline of how calculations are run in GAMESS: HF/DFT/GVB/MCSCF $\Rightarrow$ CI/MP2/CC (correlation corrections) $\Rightarrow$ SOC/Douglas-Kroll (relativistic corrections) [68, 69, 70]. The GAMESS input documentation provides further characterization and explanation⁴.

⁴https://www.msg.chem.iastate.edu/gameess/GAMESS_Manual/docs-input.txt

2.5.2 nlogn, ComputeCanada (Computational Resources)

All computations were performed on *nlogn* and *ComputeCanada* clusters. *nlogn* is a cluster available to the theoretical chemistry groups at University of Waterloo. It features Intel Westmere X5670 CPUS with 12 cores and 48 GB of memory on 6 compute nodes. ComputeCanada (Digital Research Alliance of Canada) is the national high-performance computing platform. The available computational resources are available to the Alliance wiki⁵. To take advantage of the multiple CPUs available on both platforms, the instructions and knowledge for parallelized runs in both GAMESS and MCTDH/VECC are available in Appendix A.1.

⁵https://docs.alliancecan.ca/wiki/Technical_documentation

Chapter 3

Diabatization Code Development

The newly developed software package will be described, along with technical details on how it is used to produce vibronic models. The majority of the development process was done in collaboration with Neil Raymond, and the software is publicly available on [GitHub](#) [71]. `dist_allmodes_pm.py` (`dist`) is the main diabatization code program. As previously outlined in Section 1.3, this is my main contribution, and the centerpiece of the project.

The diabatization scheme covered in Section 2.1.2 will be tackled in two broad stages: first, we employ a MCSCF-style method to obtain a ‘template’ (primarily the diabatic states, as well as a reference geometry model, frequencies and other assorted parameters) and then we proceed to express the Hamiltonian in those diabatic states, (i.e., obtain a Hamiltonian in the diabatic representation). I briefly outline in Section 3.1 the individual steps of the MCSCF calculations, and `DiabToolkit` which assists in tackling this process. Next, Section 3.2 explains how the `dist` code performs the diabatization, by parsing through an example ($\text{Fe}(\text{CO})_5$) step-by-step. In addition, I describe in detail the practical challenges encountered, as well as their solutions, in automating this extensive and unwieldy process. Workflows and software related to spectroscopy are discussed in Section 4.1. Finally, some implementation details and other concerns are briefly touched upon in Section 3.3.

3.1 MCSCF Pipeline

The MCSCF process can be broken down into four individual steps. First, a geometry optimization and frequency (Hessian) calculation is performed (A.1.1). Second, a GMC-QDPT single point calculation is conducted (A.1.2). Third, the adiabatic states and diabatic molecular orbitals (DMO) are calculated (A.1.3). Finally, a reference determinants (REFDET) calculation will return diabatic states and energies in the fourth step (A.1.4). Some steps require very specialized knowledge of GMC-QDPT and CASSCF intricacies¹. The approach taken here is a slightly modified version as seen in Zeng’s work [21]. A visual representation of the entire Zeng computational protocol (diabatization process) is presented in Figure 3.1, with the four MCSCF steps highlighted in purple.

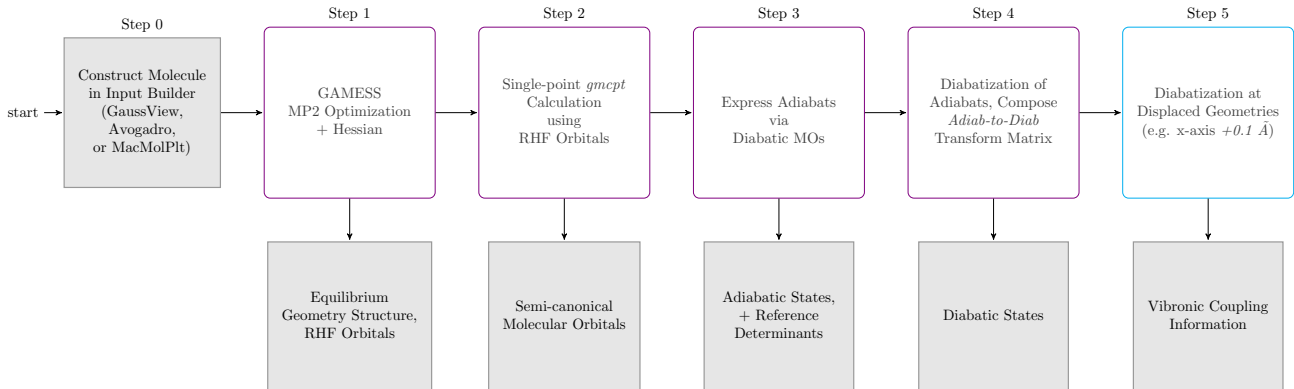


Figure 3.1: GMC-QDPT vibronic coupling model scheme formulated by Zeng.

We begin by using standard computational chemistry software (e.g., ORCA, GAMESS) to define the equilibrium coordinates and the frequencies based on the input geometry². In the second step we use a GMC-QDPT calculation to obtain the adiabatic energies and the semi-canonical molecular orbitals (of which a subset are diabatic orbitals). The CASSCF DMOs are used to calculate reference determinants in step three. Finally, in step four, $\underline{U}$ and $|\phi^d\rangle$ are calculated directly in GAMESS using all the ingredients obtained in the previous calculations (e.g., ref. geom, CMOs, REFDETs). After these four steps, all of the

¹Consultation of Nakano’s excellent student guide [72] is highly advised.

²This is essentially the Z-matrix as commonly seen in other packages.

necessary components for a diabatic model have been calculated, and the template is then used in the diabatization procedure as described below in Section 3.2. Originally, the process for performing these calculations was an assortment of manual copying, shell scripts, and excel editing tasks; which was excessively slow, error-prone, and convoluted to non-experts. Because the method is approximate and quite complex, it can be quite time consuming to distinguish between operator error and systematic properties. As a consequence the overhead of on-boarding was exceptionally high, and roadblocks encountered during execution of the method required diagnosing by a specialist.

Setting up an working active space (see Section 2.2.3) is by far the most difficult challenge [73]. This mainly focuses on the second step. Complications are primarily due to reaching a workable combo of orbitals and electrons in conjunction with preservation of orbital ranking order and orbital pureness. Another challenge is converging the Lagrangian in the MCSCF optimization sections (the bulk of the work happening during the calculation). Without intimate knowledge of various tricks of the trade such as SCF convergers and proficiency in using GAMESS, the Lagrangian calculation will often diverge. If the calculation diverges, then no energies can be obtained and it is not possible to continue with the rest of the protocol.

Other demanding elements of the MCSCF process were the composition of inputs for steps three and four. These typically required extracting information from the preceding calculation’s ‘punched-out’ orbitals and/or energies. For simple models such as H_2O , manually performing this process was burdensome, but not prohibitive. Commonly there are a couple dozen output files from GAMESS (for a triatomic system), roughly 6000 lines each, and it takes on the order of minutes to manually compose the inputs. For more sizeable systems (but still chemically small molecules, 11 atoms) such as $\text{Fe}(\text{CO})_5$, manual processing is absurd. There are 1712 output files, each of which are millions of lines long (70 MB), and in essence is ”impossible to attack” by hand.

To address these obstacles, a package of software scripts ([DiabToolkit](#)) was developed.

Although `DiabToolkit` is not the main diabaticization code we developed, it is a very useful peripheral tool. The package provides eight streamlined programs which assist with executing the protocol including such tasks as extracting information, performing *simple* mathematical calculations, preparing input files for further steps, providing helpful suggestions and warnings to the operator, and automating the validation of data within minimum specifications (choosing DMOs from CMOs, no anomalous coefficients, error checks, etc.). The `DiabToolkit` scripts have saved an incredible amount of time and frustration. They drastically reduced the on-boarding process, thereby increasing the likelihood of adoption (of this method) by new researchers. By reducing the time spent preoccupied on mechanical duties, more time will be available for innovative research. Most importantly, by automating the process, human error is almost entirely eliminated.

A detailed walkthrough of each step in the MCSCF process, using `DiabToolkit` for a RhF_3 SOC system, is provided in the appendix [A.1](#). Once Step 4 is completed, vibronic coupling information at reference geometry is directly available, and the Step 4’s input file can serve as a template for the diabaticization to proceed.

3.2 Automated Machine Gun Diabatization Code

As shown in Figure [3.1](#) the Zeng protocol can be broken down into five steps. The previous four steps (highlighted in [purple](#)) were covered, and it is prudent to note that they only touch upon the MCSCF-template pipeline component of the entire diabaticization process. Here, we expand on the final step (outlined in [blue](#)), and describe how, the `dist` code was developed to automate and address the complicated process shown in Section [2.1.2](#).

Given a template, as described in the previous section, the current `dist` code generates a vibronic model Hamiltonian (with up to quadratic order coupling constants) in a suitable manner for both VECC and MCTDH. The diabaticization of the template can be broken down into six steps ([3.2](#)). Step A, the first step, involves a selection of key parameters

in the Python project parameters file. Extraction of frequencies and undistorted reference structure coordinates, storing them inside dictionaries occurs in the second Step B (Section 3.2.1). The single point calculation at the reference geometry (refG) is submitted at Step C (Section 3.2.2). In Step D (Section 3.2.3), the fourth step, `dist` automatically performs precomputation of distortions and subsequent submission of calculations in ‘machine gun’ fashion. $\underline{U}$ is likewise calculated here. The harvesting and collection of diabatic vibronic model data is processed in Step E, (Section 3.2.4). Finally, in Step F (Section 3.2.5), at the user’s request, `dist` executes fitting of PES plots for higher order modes.

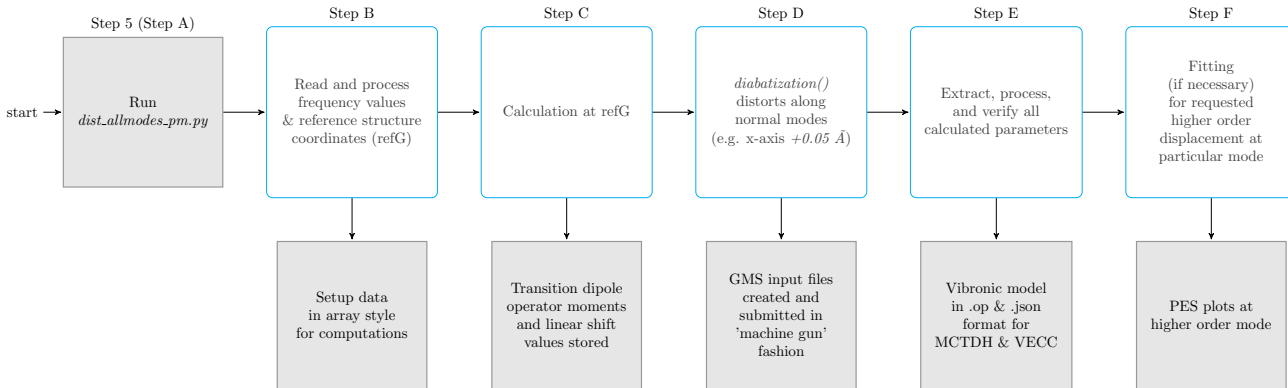


Figure 3.2: The six step process automated by the `dist` code.

In the first step (A) the user must carefully select a number of pivotal parameters including: which $3N - 6$ modes (if any) to ignore, set the number of electronic states, the magnitude of the reduced nuclear coordinate displacement (`qsize`), and the grid upon which the displacements are evaluated. For example, a ‘screened’ version of $\text{Fe}(\text{CO})_5$ might only use 12 out of the 27 vibrational modes present in the template, a `qsize` of 0.3 (unitless), and for a particularly important mode (v_{23}) use a `x8` grid³ (meaning v_{23} is evaluated at 8 points: `qsize, 2*qsize, 3*qsize, ...`). Triple checking of parameters is recommended before executing the code. Once the key parameters are chosen, the `dist` scripts automatically executes and handles the remaining five steps (B–F).

³The evaluation of the modes on a displacement grid is referred to as “high order expansions in nuclear coordinates” in Ref. [21].

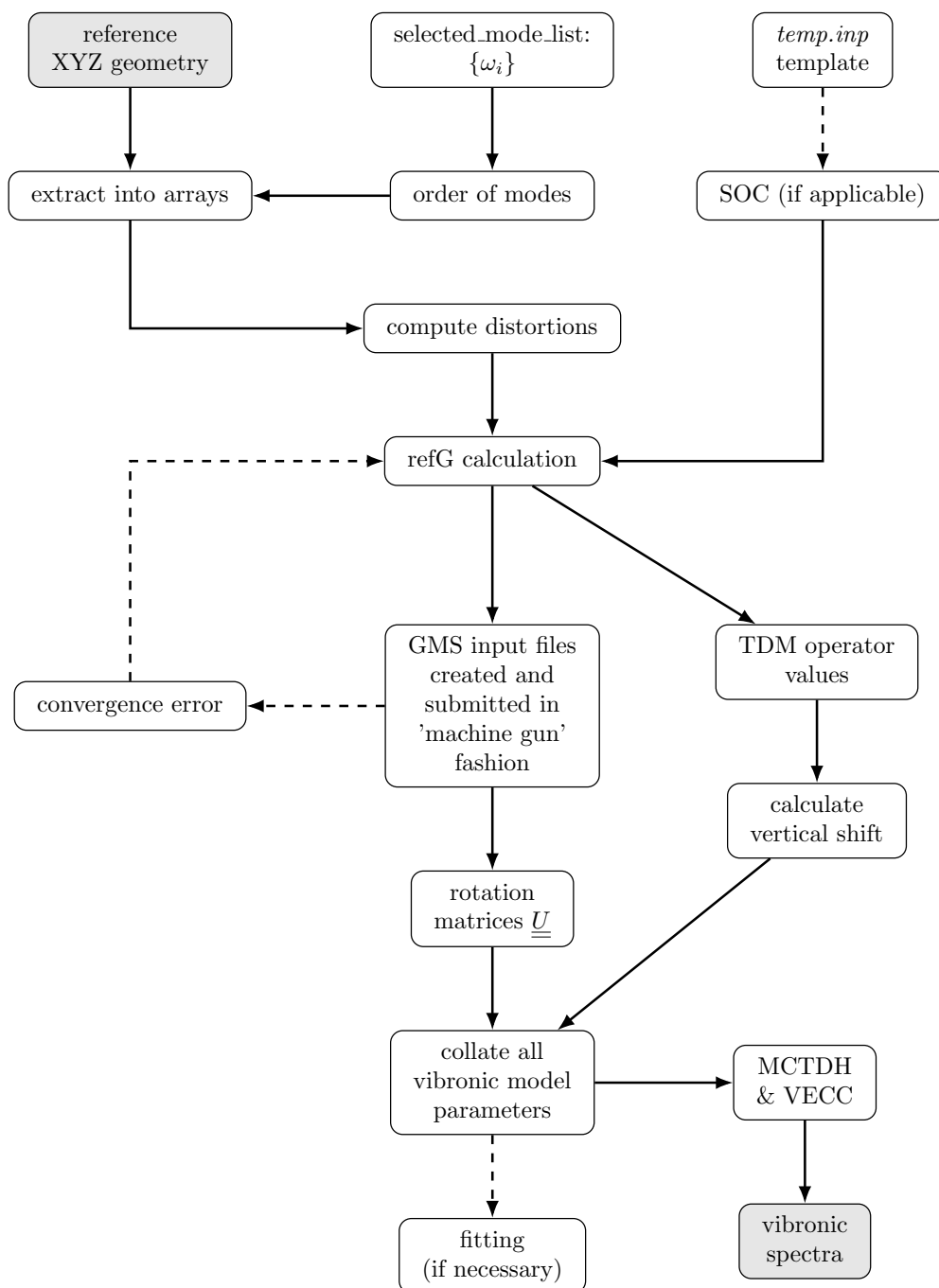


Figure 3.3: Alternative `dist` flow chart, displaying the connections and dependencies between the various steps that occur when running the program.

3.2.1 Step B: Frequencies and Reference Structure

In this step, the frequencies and reference structure’s coordinates ($\vec{R}_{q_0}$) provided by the geometry optimization and k-Hessian calculation are extracted into convenient forms. These are necessary for distorting along modes ($\vec{q}_i$, grids) and eventually obtaining new diabatic Hamiltonian matrix elements through finite difference:

$$V_{ab}^i = \frac{V_{ab}(\vec{R}_{q_0} + d\vec{q}_i) - V_{ab}(\vec{R}_{q_0} - d\vec{q}_i)}{2d\vec{q}_i}$$

as shown in Chapter 2 2.28.

Concerning geometry, **dist** starts by searching backwards from the end of the file for the first occurrence of the header **EQUILIBRIUM GEOMETRY LOCATED**, which allows us to locate the **final**⁴ block containing the reference geometry coordinates calculated during the geometry optimization; an example is provided in Listing 3.1.

109	***** EQUILIBRIUM GEOMETRY LOCATED *****				
110	COORDINATES OF ALL ATOMS ARE (ANGS)				
111	ATOM	CHARGE	X	Y	Z
112	-----				
113	FE	26.0	-0.0000000000	0.0000000000	0.0000000000
114	C	6.0	-0.8824444078	1.5284385491	0.0000000000
115	C	6.0	-0.8824444078	-1.5284385491	0.0000000000
116	C	6.0	1.7648888155	0.0000000000	0.0000000000
117	O	8.0	-1.4586915946	2.5265279545	0.0000000000
118	O	8.0	-1.4586915946	-2.5265279545	0.0000000000
119	O	8.0	2.9173831893	0.0000000000	0.0000000000
120	C	6.0	0.0000000000	0.0000000000	-1.6781093310
121	C	6.0	-0.0000000000	0.0000000000	1.6781093310
122	O	8.0	-0.0000000000	0.0000000000	-2.8433375200
123	O	8.0	-0.0000000000	0.0000000000	2.8433375200
124					
125	INTERNUCLEAR DISTANCES (ANGS.)				
126	-----				

Listing 3.1: Optimized geometry of Fe(CO)₅ coordinates at equilibrium geometry in Å.

Using **INTERNUCLEAR DISTANCES** as a footer, we then extract the xyz components of the $N = 11$ atoms (lines 113–123 in the example). When they are stored in dictionaries, we have split the line into three fields: atom, charge, and coordinates. On frequencies, we

⁴As it is an optimization/iterative method, there can be multiple instances.

similarly isolate the frequencies block (as seen in listing 3.2), and then intelligently bake the information into NumPy arrays [74].

48786	FREQUENCIES IN CM**-1, IR INTENSITIES IN DEBYE**2/AMU-ANGSTROM**2,					
48787	REDUCED MASSES IN AMU.					
48788		26	27	28	29	30
48789	FREQUENCY:	804.43	922.45	922.46	1973.40	2055.77
48790	SYMMETRY:	A2 "	E '	E '	A2 "	A1 '
48791						
48792		31	32	33		
48793	FREQUENCY:	2099.41	2248.92	2248.92		
48794	SYMMETRY:	A1 '	E '	E '		

Listing 3.2: Optimized frequencies of Fe(CO)₅.

This core regular expression (regex) pattern processing work is the bread and butter of how `dist` operates and reduces handiwork. In the prior Bash code, the reference geometry would have to be extracted by hand into a 'ref_structure' file in the same directory. Sometimes when the input symmetry of the molecule intrinsically produces a saddle point, the first frequency becomes imaginary and has an 'I' attached to its value. The code accommodates for this special case too.

3.2.2 Step C: Calculation at Reference Geometry (refG)

Step C is designed to efficiently save time and ensure the template is well-formed. It is essentially a repeat of the fourth step of the MCSCF process, where we obtained diabats at reference. If the `refG` calculation does not terminate successfully, then `dist` will exit, as proceeding is useless. Diabatization **cannot** proceed until the user has isolated and corrected any mistakes in the template⁵. In essence the output values could all be taken from the prior calculation, but this verification has been very useful in catching small and hidden errors, and this step is computationally "free".

To produce simulated spectra, this necessitated upgrading `dist` to incorporate additional spectroscopy-related components in the vibronic Hamiltonian model it returns. The original code only targeted the calculation of the vibronic Hamiltonian coefficients. We now have

⁵Explained in more detail in the Appendix.

added the capabilities of calculating vertical shifts and transition dipole moments (TDM)⁶. TDM (operator) coefficients are taken from the output refG calculation as viable TDM values as shown in Listing 3.3. Note that, theoretically, diffuse basis sets are necessary to obtain ‘true’ TDM values. VECC currently is set to use default TDM values. Once again, the workflow and repository software accounts for this distinct condition.

```

55638 TRANSITION DIPOLES BETWEEN DIRECT MAX. DIABATS (in A.U.):
55639
55640 STATE PAIR      X      Y      Z
55641    1    1 -0.000001  0.000002  0.000000
55642    2    1 -0.000000 -0.000000 -0.000000
55643    2    2 -0.019332 -0.008830  0.000000
55644    3    1 -0.000000  0.000000 -0.000000

```

Listing 3.3: TDM block as seen in the Fe(CO)₅ refG calculation.

3.2.3 Step D: Submit Diabatization Displacement Calculations

This is the critical step in `dist` which fulfills most of the steps laid out in Section 2.1.2. Here is where all the many displacement calculations required to build the grid of diabats are first calculated and then submitted to GAMESS. If we analyze 2.28, shown inside the equation are $+dq_i$ and $-dq_i$. Plus/minus distortions are precomputed on the geometry via linear, bilinear, and quadratic order displacements along the modes as seen in Listing 3.4.

```

1 displacements = {
2     "+1": reference[:, NEW] + 1.0 * R_array[NEW, :] * mode_array[:, :],
3     "-1": reference[:, NEW] - 1.0 * R_array[NEW, :] * mode_array[:, :],
4     "+2": reference[:, NEW] + 2.0 * R_array[NEW, :] * mode_array[:, :],
5     "-2": reference[:, NEW] - 2.0 * R_array[NEW, :] * mode_array[:, :],
6 }
7 #                      (Z*3, 1)                      (1, N)                      (Z*3, N)
8
9 displacements = {
10     "++": distorted_coords["+1"][:, :, NEW] + R_array[NEW, NEW, :] * mode_array[:, NEW, :],
11     "+-": distorted_coords["+1"][:, :, NEW] - R_array[NEW, NEW, :] * mode_array[:, NEW, :],
12     "-+": distorted_coords["-1"][:, :, NEW] + R_array[NEW, NEW, :] * mode_array[:, NEW, :],
13     "--": distorted_coords["-1"][:, :, NEW] - R_array[NEW, NEW, :] * mode_array[:, NEW, :],
14 #                      (Z*3, N, 1)                      (1, 1, N)                      (Z*3, 1, N)
15 }

```

Listing 3.4: Precomputing linear, quadratic, and bilinear displacements in `dist`.

⁶and also for MCTDH the introduction of a fictitious ground state (for photoexcitation spectra) but this does not occur in Step C.

The naming convention is a significant parameter when it comes to `dist` runs where there can many thousands of files in the target directory. Therefore, we have designed a clear and concise naming scheme where the prefix concatenates global parameters (symmetry, basis sets, etc) of the system and importantly, a suffix specifying `{qsize}`, `{d1}`, `{order}`, `{q1_label}` and `{d2}`, `{order}`, `{q2_label}` if bilinear. This suffix describes order of displacement along modes and is what separates each file in the directory. For example, using $\text{Fe}(\text{CO})_5$, with Sapporo basis set, C1 symmetry, 17 states, at negative linear displacements of 0.05 on modes 31 and 27, would be `FeCO_SPK_gmcpt_C1_17st_diab_0.05_-x1q31_-x1q27.inp`.

The new coordinates for the system are calculated, and then stored inside their respective input file. Input files copy the template and only modify its coordinates from the original in Listing 3.1. Specifically, for both the linear and quadratic case, the code will generate a $+/-$ calculation for each mode from 1 to n , where n is the highest order requested at a mode. On the bilinear case, the code will proceed to generate four calculations at each mode-mode pair displacement: `"++": 'pp'`, `"+-": 'pm'`, `"-+": 'mp'`, `"--": 'mm'`. Once this processing is complete, all calculations are submitted to GAMESS. However, if in the selected mode list there are higher order displacements requested at a certain mode (e.g., x8 instead of the standard x2), then additional $+/-3 \dots +/-8$ calculations will be precomputed and submitted for that mode.

1	ATOM	ATOMIC	COORDINATES (BOHR)		
2		CHARGE	X	Y	Z
3	FE	26.0	-0.0000021335	0.0012378938	0.0000000002
4	C	6.0	-1.6661297582	2.8864440831	-0.0000000004
5	C	6.0	-1.6674766864	-2.8875279147	-0.0000000026
6	C	6.0	3.3336072356	0.0006245327	0.0000000039
7	O	8.0	-2.7568757238	4.7748782428	-0.0000000001
8	O	8.0	-2.7572011762	-4.7757841321	0.0000000042
9	O	8.0	5.5140787632	-0.0001710964	-0.0000000043
10	C	6.0	0.0000049152	-0.0028511360	-3.1707512585
11	C	6.0	0.0000049197	-0.0028511336	3.1707512453
12	O	8.0	-0.0000011801	0.0006848022	-5.3734056464
13	O	8.0	-0.0000011785	0.0006847984	5.3734056520

Listing 3.5: -x1q31_-x1q27 distorted coordinates (negative 31st and 27th mode linear order).

The Diabatic Rotation Matrix $\underline{\underline{U}}$

At each diabaticization displacement calculation submitted by `dist`, the crucial diabatic rotation matrix $\underline{\underline{U}}(\vec{R}_{\pm dq})$ is calculated by GAMESS. Whenever the condition that both the DMO and REFDET groups are present (in the input file) is fulfilled, the $\underline{\underline{U}}$ matrix for the GMC-QDPT states will be calculated as shown in Listing 3.6. The number of diabats is consistent to the number of adiabats, and will maintain the phase as given in the REFDET group. If only DMO group is present, in this occasion, it will only yield the transformation matrix for CMO to DMO and adiabatic states expressed in terms of DMOs.

52836	-----					
52837	- ROTATION TO DIABATIC STATES -					
52838	-----					
52839	ROWS=ADIABATIC , COLUMNS=DIABATIC STATES					
52840						
52841		1	2	3	4	5
52842						
52843	1	-1.0000000	-0.0000000	0.0000000	0.0000000	0.0000000
52844	2	-0.0000000	-0.0530665	0.1817172	0.3081071	-0.0167091
52845	3	-0.0000000	0.7366603	-0.2748155	-0.4347364	0.1508355
52846						
52847	17	-0.0000146	-0.0000000	0.0000000	0.0000000	0.0000000
52848						
52849		6	7	8	9	10
52850	1					

Listing 3.6: Fe(CO)₅ 17 states rotation matrix U calculated within GAMESS.

An Improved Recipe for Distortions

Loops in python are notoriously slow [75]. We initially emulated the original Bash prototype implementation of the distortion calculations using loops. So we refactored the excessive looping that distorted the modes into efficient vectorized array operations shown in Listing 3.4. This drastically improved the performance. Furthermore, it also made the entire program more dynamic in a sense that it can adapt and be extended for a variety of purposes. An example here is fitting. This situation arose where distortion beyond-quadratic-order at a single mode was needed, and the recasted arrays handled this trivially as compared to looping. Another key aspect of this change was the fact that the arrays forced us to understand exactly what dimensions are required in the diabaticization code's various calculation modules.

Without this directed and compelled learning, much of the pedagogical value in the exercise of reformulating a new diabaticization code and 'learning the ropes' would not have occurred. Circumstantially, the code would also be unwieldy and resistant to accommodating changes.

3.2.4 Step E: Harvest Results for MCTDH/VECC

As per standard MCTDH Hamiltonian format, the operator file containing the so-called model Hamiltonian is constructed procedurally in various sections. The first section here is the operator-section, which initializes the title of the MCTDH run. Next, the frequencies, state energies (diagonal/off-diagonal pair), and electronic TDM values are defined.

```

1 operator-section
2 frequencies
3 E0_energies (diag/off-diag pair)
4 electronic TDM
5 parameter-section
6 linear coupling
7 bilinear coupling
8 quadratic coupling
9 {SOC constant} (if necessary)
10 {SOC linear} (if necessary)
11 {SOC bilienar} (if necessary)
12 {SOC quadratic} (if necessary)
13 hamiltonian-section
14 hamiltonain-section_operator

```

Listing 3.7: MCTDH operator file format constructed with standard ingredients in a vibronic model.

Inside the parameter section are the linear, bilinear, and quadratic coupling term values. If the calculation is SOC enabled, then SOC version of the couplings will similarly be defined. This includes constant terms, which exist as E0_energies prior to the parameter-section. Diabatically, these energies appear in the GAMESS output as 'GMC-PT-LEVEL DIABATIC ENERGY=', and critically, if they are diagonal in the states array ([a,a]) then it is extracted in Hartree units. If they are off-diagonal ([b,a]) diabatic couplings, then they will be extracted in eV since at displaced geometries, the printed values are relative to the reference geometry. The regular expression (regex) is slightly different, as the printed value will be shown as 'GMC-PT-LEVEL COUPLING' instead, as indicated in Listing 3.9. Moreover, the extracted

off-diagonal values will fill the upper-triangular part of the matrix with rows and columns defined by states.

```

1 extracted_Hartree_1D = extract_in_Hartrees(path, a_pattern.format(col=0+1), cheat=True)
2 for a in range(A):
3     array[a, a] = extracted_Hartree_1D[a]
4
5 extracted_eV_2D = extract_in_eV(path, ba_pattern.format(row=0+1, col=1+1), cheat=True)
6 for a, b in upper_triangle_loop_indices(A, 2):
7     array[a, b] = extracted_eV_2D[a, b]

```

Listing 3.8: Differentiation of linear vibronic coupling values extraction in the dist program.

```

1 -----
2 - DIABATIC HAMILTONIAN MATRIX ELEMENTS -
3 -----
4
5 --- DIABATIC ENERGIES (DIAGONAL ELEMENT) ---      HARTREE      EV
6 STATE # 1'S GMC-CI-LEVEL DIABATIC ENERGY=      -1835.393289739      0.000000000
7 STATE # 2'S GMC-CI-LEVEL DIABATIC ENERGY=      -1835.211881140      4.936379413
8 .....
9 STATE # 16'S GMC-CI-LEVEL DIABATIC ENERGY=      -1835.157909542      6.405021408
10 STATE # 17'S GMC-CI-LEVEL DIABATIC ENERGY=      -1835.157909551      6.405021138
11 .....
12 --- DIABATIC COUPLINGS (OFF DIAGONAL ELEMENTS)---      HATREE      EV
13 STATE # 1 & 2'S GMC-CI-LEVEL COUPLING =          0.000000000      0.000000000
14 STATE # 1 & 3'S GMC-CI-LEVEL COUPLING =          0.000000000      0.000000000
15 .....
16 STATE # 15 & 17'S GMC-CI-LEVEL COUPLING =          0.000000004      0.000000105
17 STATE # 16 & 17'S GMC-CI-LEVEL COUPLING =         -0.000000004     -0.000000110

```

Listing 3.9: Diabatic matrix elements example found in GAMESS.

Linear Coupling Terms

The linear coupling terms are calculated as shown in Equation (3.1). The temporary dictionaries (*temp_dict*) found in Listing 3.10 are the data structures used to temporarily hold the coupling term values extracted from GAMESS output files to be used for computation.

$$H_{\mu\lambda, Linear}(q) = \sum_i^M E_{\mu\lambda}^i q_i ; E_{\mu\lambda}^i = \frac{E_{\mu\lambda}(\vec{R}_{q_0} + dq_i) - E_{\mu\lambda}(\vec{R}_{q_0} - dq_i)}{2dq_i} \quad (3.1)$$

```

1 linear_ev = (temp_dict["+1"] - temp_dict["-1"]) / (2*qsiz)
2 for a in range(A): # make sure we multiply the diagonal by ha2ev
3     linear_ev[a, a] *= ha2ev

```

Listing 3.10: Evaluating linear coupling term and conversion of units. Not the full code implementation in `dist`, demonstration purposes only.

Quadratic Coupling Terms

In the quadratic terms, we copy over the entire +2/-2 arrays. The code shown in Listing 3.11 implicitly adds the entire arrays. Only the diagonal is multiplied by conversion factor *ha2ev* (omitted in Listing 3.11).

$$H_{\mu\lambda, Quadratic}(q) = E_{\mu\lambda}(0) + \frac{1}{2} \sum_{i=j}^M E_{\mu\lambda}^{ij} q_{i=j}^2 \quad (3.2)$$

```

1  quad_ev = temp_dict["+2"] + temp_dict["-2"] - 2.0 * E0_array
2
3  quad_ev /= (2. * qsize) ** 2
4  for a in range(A):
5      quad_ev[a, a] -= vibron_ev[i]
```

Listing 3.11: Evaluating quadratic coupling term and conversion of units. Not the full code implementation in `dist`, demonstration purposes only.

Bilinear Coupling Terms

Bilinears calculations produce a set of four terms, and are characteristically denoted by pp, pm, mp, mm.

$$H_{\mu\lambda, Bilinear}(q) = \sum_{i,j;i>j}^M E_{\mu\lambda}^{ij} q_i q_j \quad (3.3)$$

```

1  bilin_ev += temp_dict['++'] - temp_dict['+-']
2  bilin_ev += temp_dict['--'] - temp_dict['-+']
3  for a in range(A): # make sure we multiply the diagonal by ha2ev
4      bilin_ev[a, a] *= ha2ev
5  bilin_ev /= (2. * qsize) ** 2
```

Listing 3.12: Evaluating bilinear coupling term and conversion of units. Not the full code implementation in `dist`, demonstration purposes only.

After turning on SOC in the second step of the MCSCF process, diabatic spin orbit matrix elements are available. This is an addition to the regular ordinary diabatic non-spin orbit matrix elements (DSOME). To compute the diabatic SOC matrix elements, it

From usage in spectroscopy, it was revealed that there was a need for truncated model Hamiltonians. For $\text{Fe}(\text{CO})_5$, the full length operator file for 17 states contains a staggering 639,421 lines! However, by truncating the model and omitting zero-valued couplings, the file size is reduced to approximately 260,000 lines—which is still formidable, but more manageable. Since the code was compartmentalized and systematically built in sections, this was a feasible change. Now, `dist` will automatically perform all the correct symmetrizations and then output the truncated constant and linear only versions of the operator file. For easy comparison with VECC, the same model in operator file format is converted to JSON format. The code actually brute-force calculates all vibronic coupling terms, even if they are known to be zero by symmetry. However, it will know to suppress zeroed terms in printing the results of the model. At this point, the code has prepared all the ingredients needed from the GAMESS calculation for dynamics: diabatic state energies, frequencies, electronic TDMs, vibronic couplings, and (if applicable) SOC couplings.

3.2.5 Step F: Fitting on n-th Order Models (Conditional)

We have written a module in `dist` to take care of fitting. Similar to the PES depictions found in Chapter 2, there was a need for drawing PES plots from the diabatization data. Fitting became highly relevant when we applied our code to study both D_{3h} tetratomic molecules and $\text{Fe}(\text{CO})_5$, as shown in Figure 3.5 and Figure 3.6. Additionally, to test the correctness of our new array-style computation of distorted coordinates, we verified through fitting the diabatic state energies. If the user requested (beyond second, > 3) higher order multiplication of qsize displacement at a mode, then `dist` will proceed to perform fitting automatically. Importantly, GAMESS’ output differs depending on if the calculation is single or multi-surface. `dist`’s fitting knows to differentiate extracting a single or multi-surface scenario as it will search for `E(GMC-PT2)` and `E(GMC-QDPT2)` values respectively. Once all fitting values are obtained, `dist` will link together the energies using gnuplot’s least-squares `fit` algorithm and plot it.

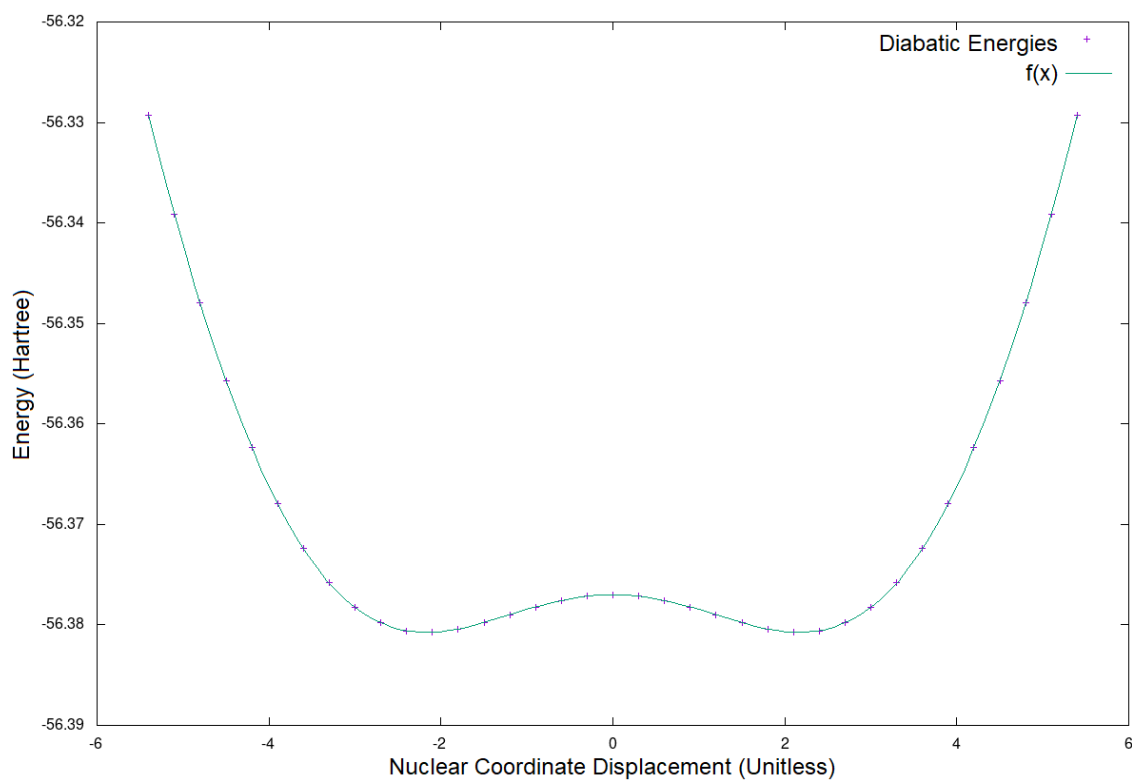


Figure 3.5: PES of D_{3h} NH_3 , Mode 7, ± 10 Displacements, reduced nuclear coordinate displacement ($qsize$) = 0.30 per point.

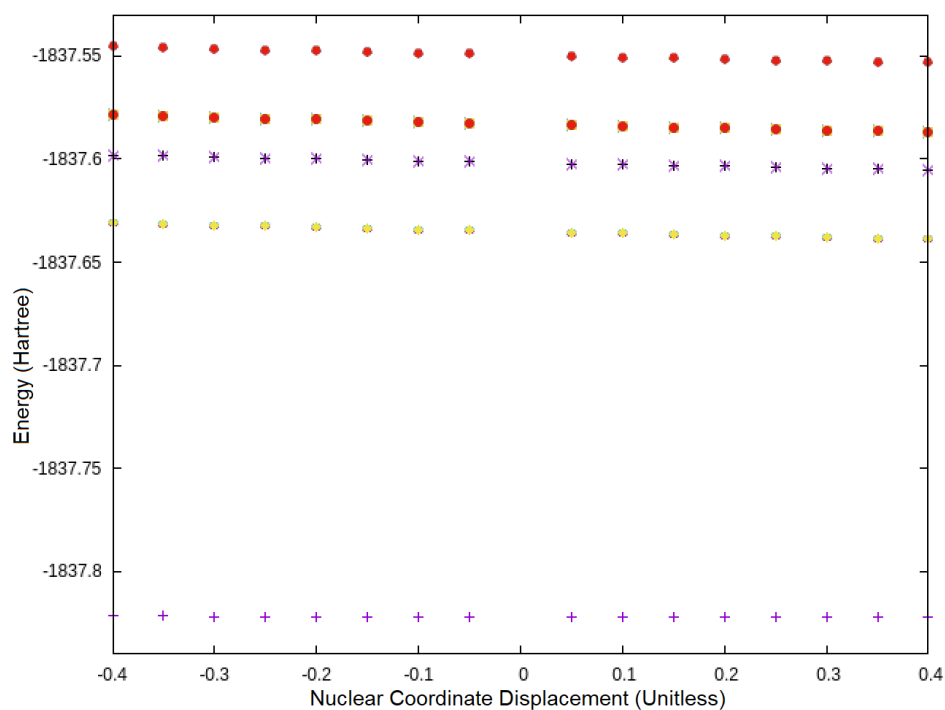


Figure 3.6: PES of $\text{Fe}(\text{CO})_5$, composed from E(GMC-QDPT) single point diabatic energies at mode 15, ± 8 Displacements, reduced nuclear coordinate displacement ($qsize$) = 0.05 per point. Both plots are created from `dist`'s fitting module.

3.3 Assorted Comments and Considerations

3.3.1 Useful Regular Expressions

On the *nlogn* cluster, due to an obscure bug in the compiled GMCPT-GAMESS version, whenever SOC calculations are run, the last approximately 200 lines of the output file will **always** show random corrupted bytes of memory such as: `Error termination. Backtrace: ##0 0x7f2b79702d21 in ??? ##1 0x7f2b79703869 in ??? ...` and so on. Therefore, it was determined a posteriori that 'DONE WITH MP2' is greatest working footer (regular expression) keyword in every single situation, even memory-corrupted SOC calculations. On ComputeCanada clusters, SOC calculations have no issues.

3.3.2 Memory Mapping

Memory mapping is a technique where a file is directly mapped into a computer's virtual memory and can be read highly efficiently [76]. We originally opted to extract data using the slow pythonic 'read' method. However, this proved to be debilitating when operating on large output files to a point where the bottleneck could be days when running the script. Therefore we opted to use the `grep` and `sed` command instead. When it came down to the full computational study on $\text{Fe}(\text{CO})_5$, surprisingly even `grep` was deemed too slow⁷. The average size of the $\text{Fe}(\text{CO})_5$ output file is 70 megabytes or around 2 million lines of text. When we grepped the SOC data, it took hours. Since `grep` is a command line method, it depended on the speed of server calls. So to accomplish the intrinsic need for speed up over grepping, finally we shifted to using memory-mapping. This not only made the program more performant, reducing the runtime of the script processing all 27 modes of FeCO from 10300 seconds (with `grep`) to a mere 3870 seconds. However, it should be noted that we did not entirely reject `grep`, as there are still cases where a simple value lookup opts for `grep`.

⁷It is technically possible to achieve similar results using `grep` and other command line instructions, but the required high level of proficiency was not possible given project constraints and scope.

3.3.3 Effective Vibronic Coupling Magnitude: Screening

We have written a function that takes care of the effective (linear) vibronic coupling magnitude Equation (3.4). This can be referred to as screening. Please note the different meanings of "screening" here and in Section 2.3.1. The origin of this calculation is sourced from Ref. [22]. Δ_{IJ} is a difference between the diagonal matrix elements at reference structure at a particular vibrational mode, and $V_{IJ,i}^{(1)}$ is a linear vibronic coupling term. While this does not target second order and bilinear couplings (there is a different formula), linear couplings often suffice for the assessment. The premise behind this screening is to detect the strong versus weak couplings in the model and select the most dominant modes. Users can request screening by turning a switch online in the parameters. This will compare the calculated coupling values to a preset (default is -0.60) screening value and delete them from the model Hamiltonian if they fall below the value.

$$\log_{10} \left\| \frac{V_{IJ,i}^{(1)}}{\Delta_{IJ}} \right\| \quad (3.4)$$

3.3.4 Computational Roadblocks (*nlogn* and MCTDH)

Due to the immense size of the $\text{Fe}(\text{CO})_5$ system, with its full 27 modes, it became apparent that University of Waterloo's *nlogn* cluster did not have enough computational resources to handle the task. The first point of difficulty arose when there simply was not enough computational nodes on *nlogn* to perform all diabaticization calculations in a timely manner. If the full 27 modes necessitated 1712 calculations to be performed, *nlogn* is estimated to be fully consumed for more than an entire academic term. This was deemed computationally inconvenient and infeasible. Therefore, we had to switch to ComputeCanada clusters to remedy the issues, which allowed a huge speedup for us and helped finish all $\text{Fe}(\text{CO})_5$ single point calculations within a month. For a glimpse into the amount of compute needed to achieve the $\text{Fe}(\text{CO})_5$ vibronic model, ComputeCanada has recorded the quantity of CPU usage. If the average research group is allocated a default of 300 core years, then the $\text{Fe}(\text{CO})_5$

model consumed an astonishing quarter (74.33) of the core years as shown in Figure 3.7.

Although the diabatisation calculations were finished, another setback needed to be addressed: the MCTDH84 version already compiled on nlogn could not handle the maximum parameter size needed for the propagation run. There were simply too many degrees of freedom. After conversing with Prof. Dieter Meyer of MCTDH, the solution was to upgrade to the latest (2023) version of MCTDH86, which has new features for users doing large systems in particular. An ALLOC-SECTION was introduced, which allows for users to directly and dynamically request for a max parameter size without recompilation. Prof. Dieter Meyer also explained that there was also better memory management, where in the 84 version the stack had a maximal limit, now in 86 the stacks are no longer used and instead "only your computers memory is a hard limit to their sizes". Upon upgrading to 86, the propagation runs could now be performed to their full extent.

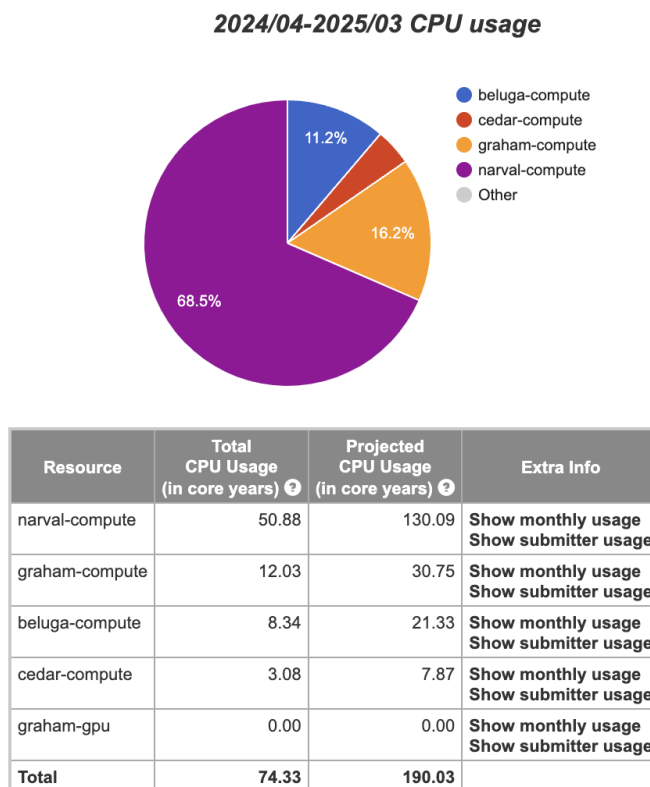


Figure 3.7: ComputeCanada usage breakdown during the 2024 academic year. The entirety of the second column CPU usage was purely dedicated to $\text{Fe}(\text{CO})_5$ diabatisation calculations.

In Summary

We have stitched together various functionalities such as: text processing, coordinate distortion, input file generation, GAMESS job submission, operator file generation, and even fitting plots into **dist**. Many bottlenecks and obstacles were overcome. Our augmentations have elevated Zeng’s prototype Bash code into an efficient highway for the automatic generation of vibronic Hamiltonians. We are confident that the sum of these parts have ultimately enriched **dist** into an extensible diabaticization code that can be useful as a catalyst to future vibronic studies.

Chapter 4

Realistic Applications to Systems of Interest

The results and analysis found in this chapter remarks on the molecular vibronic spectra produced throughout this project. After the diabaticization code generated vibronic models from the Zeng protocol GMC-QDPT(-SOC) calculations, the validity of the diabatic energies and vibronic coupling constants were tested through spectroscopic simulations. Alongside the diabaticization code development in the preceding chapter, parameterization of the project’s spectroscopy component was pursued in parallel, as detailed in Section 4.1.

We examined test-case models (H_2O , NH_3 , and PH_3) in Section 4.2 to establish initial baseline parameters for photoelectron spectroscopy. The primary challenge in parameterization was sorting out the complexity and difficulties suffered from high dimensionality. Through rigorous testing and fine-tuning, we resolved many tedious issues that caused serious friction and ultimately we established a compatible workflow that connects the vibronic model stored inside the operator/JSON file to Raymond’s in-house spectra scripts [77].

The spectra simulations shown here are highly sensitive to the many input model parameters, which heavily influence their overall shape and appearance. That is to say, they are

highly influenced by whichever input model parameters are given and tuned. The interplay of the parameters sometimes even causes unphysical results (negative intensity) or incompatibilities. Critical factors include symmetry of the original system, ionic or neutral character, and inclusion of quadratic and higher order terms. However, the aim of the research is not to reproduce “*beautiful*” (i.e. highly resolved), nor experimentally accurate spectra. Instead, the sole focus is to demonstrate the vibronic model process and its realization of simulating vibronic spin-orbit coupling effects systematically.

All spectra produced in this thesis are assumed to be of gas-phase molecules. The preliminary systems are all UV photoelectron spectra, while transition metal compound spectra pertain to UV absorption spectra of the neutral molecule. To interpret the legend in spectra, Hn represents Hamiltonian truncation order (i.e. $H0$ is constant, $H2$ is full quadratic), and Zn is the Z truncation order in VECC ($Z3$ is highest). After formulating generalized compatible parameters for photoelectron simulations using preliminary systems, in Section 4.3 we moved onto applying our methods to more sophisticated transition metal complexes. Starting with D_{3h} trigonal planar CoF_3 and RhF_3 species, since all the chemical systems listed so far were only tetratomic at best, we ultimately extended our approach to a larger $\text{Fe}(\text{CO})_5$ that would culminate as the most realistic application of the protocol to date.

4.1 Generating Spectra via MCTDH and VECC

4.1.1 MCTDH Propagations

To generate spectra in both MCTDH and VECC alike, time auto-correlation functions (ACF) are computed by propagation, and then fed into the MCTDH `autospec86` analysis program to Fourier transform and generate plotting data [65]. Here, we used a constant value of $\tau = 40$ for all runs¹. Much of the output format for `dist` was developed to seamlessly connect with

¹ τ is a weight parameter that determines time decay in autospec. It was found to be the most compatible common value to all runs.

Raymond’s project scripts, and in a mutable style based on model parameters. Formerly, the spectra scripts were configured to be able to run batch testing of different combinations of parameters on a given vibronic model; which proved useful in this project when testing newly developed GAMESS vibronic models. In MCTDH, in addition to the operator file, an input propagation (.inp) template file needs to be given to the program. This particular file will define the femtosecond propagation time, TDH or MCTDH², harmonic oscillators, basis size, maximum allocation of factors, and associated parameters. To deem whether or not a MCTDH propagation as depicted in Figure 4.1 is accurate, it is recommended that one scrutinizes the *rdgpop86* and *rdcheck86* programs, as advised by the MCTDH documentation [65]. Typically, if the spectra contains negative magnitude in spectra, it indicates too high of a τ value.

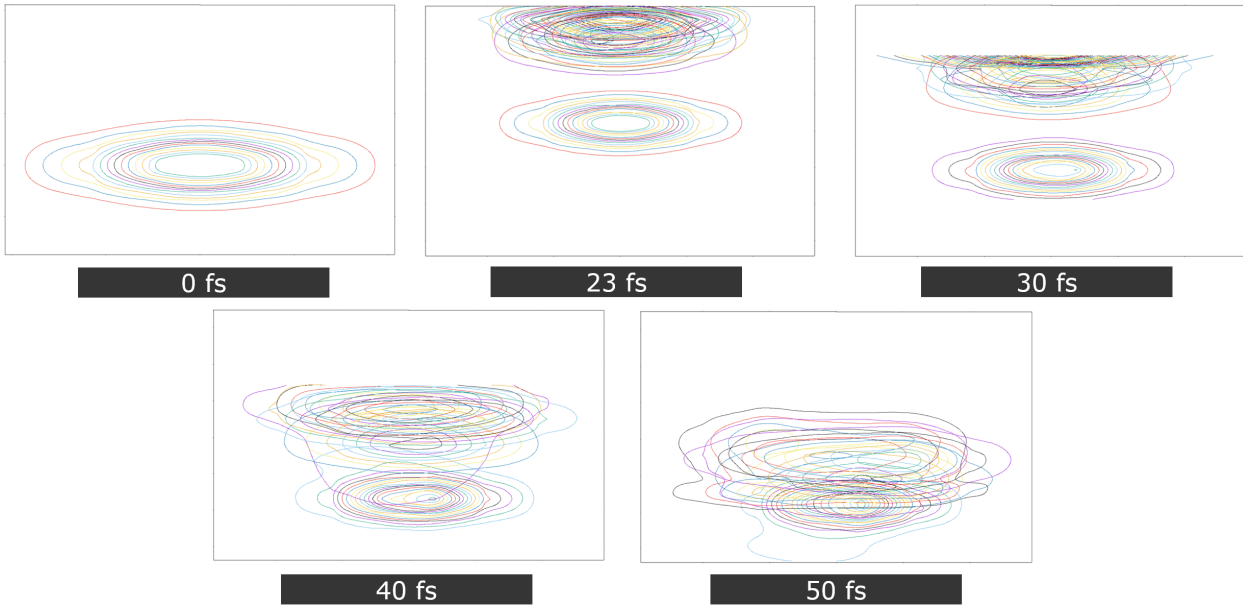


Figure 4.1: Snapshots of a $\text{Fe}(\text{CO})_5$ 17 states MCTDH 50 fs propagation. The wavepacket density excites and then proceeds to relax.

In our MCTDH simulations, we deliberately create a so-called fictitious ground state (S^{00}). This fictitious ground state is where we vertically excite the wavepacket from. Our implementation automatically accounts for this state.

²Couplings between different d.o.f. can be specified here.

4.1.2 SOC: Imaginary Changes Everything

The capacity to draw SOC-augmented spectra (correctly) was not easy to incorporate into the workflow. To fulfill extraction of DSOME mentioned earlier, this required us to reformat our code to accept a `SOC_flag`, as it is useful in instances to clearly delineate where extra SOC processing work needs to be performed and where it could be skipped. It was also a prime reason that stirred us into looking into memory mapping (Section 3.3.2), as SOC systems were starting to take significant time for `dist` to process. SOC also delivered another lesson to us: ensuring Hermiticity in the model Hamiltonians we produce. Without this clarity from SOC, the code in both `dist` and VECC would have never accommodated for this case³ and simply copy the upper triangle to the lower triangle symmetrically no matter the calculation.

Originally, VECC was implemented such that the input model it takes was strictly real-valued matrices. Yet another caveat of VECC was that it only accepted models with diagonal vertical energies (prior to this thesis project). SOC therefore became an inflection point. It motivated us to modify and adapt VECC so that it can take complex matrices and off-diagonal electronic energies. (Reminder: Section 2.3.1 SOC's $\hat{t}$ brings an imaginary component). For a long time, we were preoccupied with rationalizing why a mysterious peak appeared in the SOC spectra of MCTDH, whereas VECC had none. After an intense investigation into the discrepancy, a key fix was revealed to us whereby in MCTDH SOC propagations, if the user requests the default *auto* keyword, MCTDH actually: “assumes a real initial state and a symmetric Hamiltonian” [65]. This is incorrect for SOC runs, which are complex-valued. Therefore, the solution here was to use the *cross* keyword instead. Remarkably, this change resolved all the discrepancy issues. Every MCTDH spectra with SOC in this thesis consequently uses cross-correlation functions. This further highlights the delicateness and fragility of simulating spectroscopy. The fix was not apparent or clear cut to us at all; we spent much time and effort mired in aligning the spectra.

³Also by association, the use of complex conjugations.

4.1.3 VECC Propagations

For VECC propagations, they will similarly generate a normalized ACF file. The important aspects of VECC to configure is the balance between Z-order and propagation time. For many small models, VECC has a tendency to diverge when quadratic order coupling terms are given. This is also the case if one requests too long of a time⁴. Divergence here means that the ACF amplitude goes to infinity (typically normalized to be bound between zero and one). As mentioned in Section 2.4.2, the $\hat{Z}_x$ truncation order (of the infinite expansion) enforces the accuracy of the calculation in VECC. To date, we have not developed full intuition to ascertain whether a propagation will diverge or not prior to running it. As the Nooijen group claimed in Ref. [9], VECC is highly optimized, i.e., it performs the dynamics in a fraction of the time as compared to MCTDH. Since VECC is also implemented in Python, it is very easy to customize and adapt (simple flags, clear areas to enter parameters, etc.) towards ad-hoc situations as they require. Although VECC is less systematic and robust, we want to reiterate its advantage of being more user-friendly. The steps to draw the spectra are identical to MCTDH using its `autospec86`, and will not be repeated here.

4.2 Preliminary Systems

Here I show results for three relatively simple and small systems: H_2O , NH_3 , and PH_3 . The purpose of these simple test-case organic molecules is twofold: pedagogical examples used in helping develop our diabatization code, and to demonstrate that the vibronic model protocol has few errors in easy cases. All preliminary system species were stable enough to perform full valence active space simulations. The results will be benchmarked to He 584 Angstrom photoelectron spectra. In the appendix, there is an extra CH_2O result.

⁴You can salvage the run with high `autospec iexp` damping, but it is fickle.

4.2.1 Water, H₂O

Water was the first molecular species investigated. In this research project, the first batch of pedagogical GMC-QDPT calculations were performed on water. The parameters for the H₂O vibronic model are tabulated in Table 4.1. This model has 3 electronic surfaces ($A = 3$) and 3 normal modes ($N = 3$). $S^{II,I=1to3}$ gives the vertical ionization energies from the ground state of the neutral molecule, whose zero-point-energy-corrected energy is given by S^{00} . $C1_i^{II}$ ($C1_i^{I \neq J}$) is the linear diagonal (off-diagonal) vibronic coupling constant of diabat I (between Diabats I and J) along mode i . Whereas $C1b$ and $C2$ denote bilinear and quadratic respectively. The active space of H₂O is 6o7e. The parameters used in the GAMESS calculation are listed in Listing 4.1.

```

1      nmofzc=0 nmodoc=1 nmoact=6 nelact=7 mstart(1)=2 icharg=1
2      mult=2 gbasis=ccd state(1)=3

```

Listing 4.1: Pedagogical example of water active space parameters in GAMESS input file.

The water photoelectron spectrum is characterized by three prominent well-resolved electronic transition bands in the computed spectrum. This is conventionally reported as the $\tilde{X}^2B_1$, $\tilde{A}^2A_1$, and $\tilde{B}^2B_2$ state bands in previous photoelectron studies [78, 79]. Water in this calculation exhibits no degenerate or near-degenerate electronic states. In comparison to the experimental water photoelectron spectrum taken by Brundle and Turner [78], the simulated spectra appears to have reproduced it well and in good agreement. Evidently, the second electronically excited observed transition, induced the dominant transition centered around 17-21 eV. This band appeared to have the most discrepancy between MCTDH/VECC simulations when examined in Figure 4.3. Propagation runtimes for water are typically very fast, often finishing within minutes. This is expected to be the case for all further species with atom counts below 6 or restricted degrees of freedom. Since this was the case, a deliberately high propagation simulation time of a picosecond was requested for maximal accuracy.

Table 4.1: H₂O model computed parameters. All numerical values are in eV. The model includes constant, linear, bilinear, and up to quadratic order vibronic coupling terms.

Parameter		Linears		Quadratics	
S^{11}	11.711 685	$C1_1^{11}$	0.072 501	$C2_{11}^{11}$	-0.048 285
S^{22}	14.154 771	$C1_2^{11}$	-0.298 327	$C2_{22}^{11}$	0.022 432
S^{33}	18.116 210	$C1_1^{22}$	0.668 782	$C2_{33}^{11}$	-0.000 133
S^{00} (Z.P.E)	0.588 963	$C1_2^{22}$	-0.319 996	$C2_{11}^{22}$	-0.031 868
ω_1	0.208 018	$C1_1^{33}$	-0.728 601	$C2_{22}^{22}$	0.020 133
ω_2	0.481 209	$C1_2^{33}$	-0.835 567	$C2_{33}^{22}$	0.006 789
ω_3	0.494 264	$C1_3^{23}$	-0.234 174	$C2_{11}^{33}$	-0.223 050
				$C2_{22}^{33}$	0.081 650
				$C2_{33}^{33}$	0.037 978
		Bilinears			
		$C1b_{12}^{11}$	0.007 240	$C2_{11}^{13}$	0.000 002
		$C1b_{12}^{22}$	-0.044 403	$C2_{22}^{23}$	0.000 002
		$C1b_{12}^{33}$	0.138 827		
		$C1b_{13}^{12}$	-0.000 002		
		$C1b_{12}^{23}$	-0.000 002		
		$C1b_{13}^{23}$	-0.010 482		
		$C1b_{23}^{23}$	0.008 958		

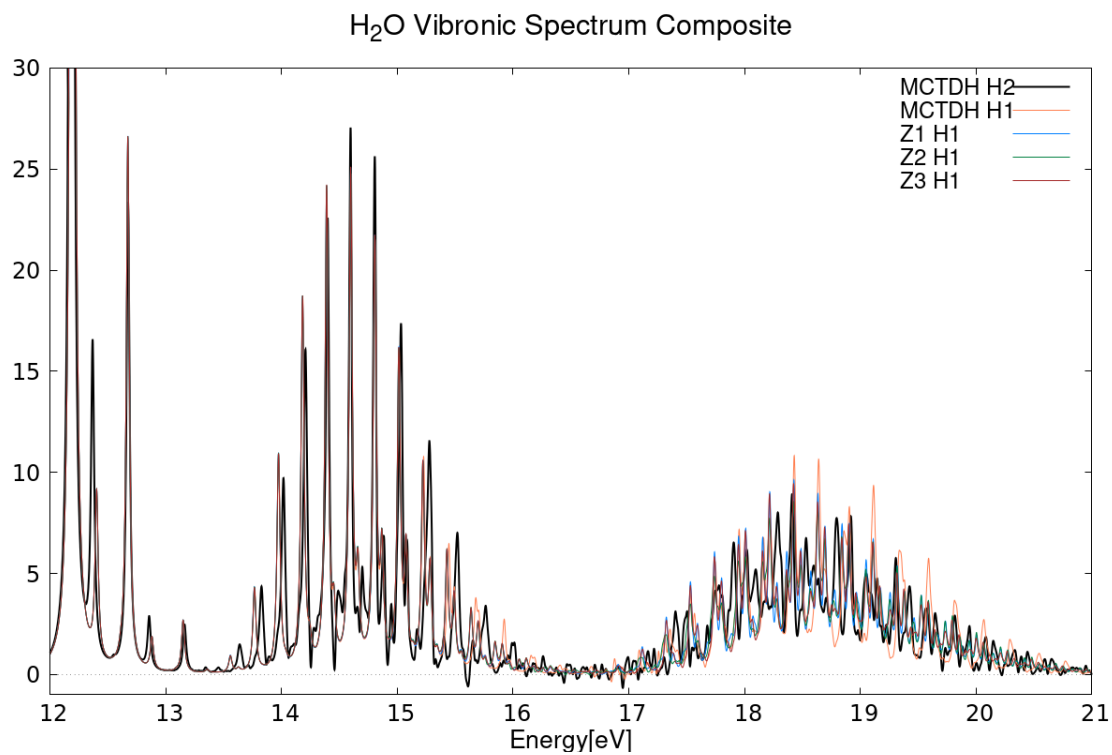


Figure 4.2: Theoretical vibrationally-resolved photoelectron spectrum of H_2O .

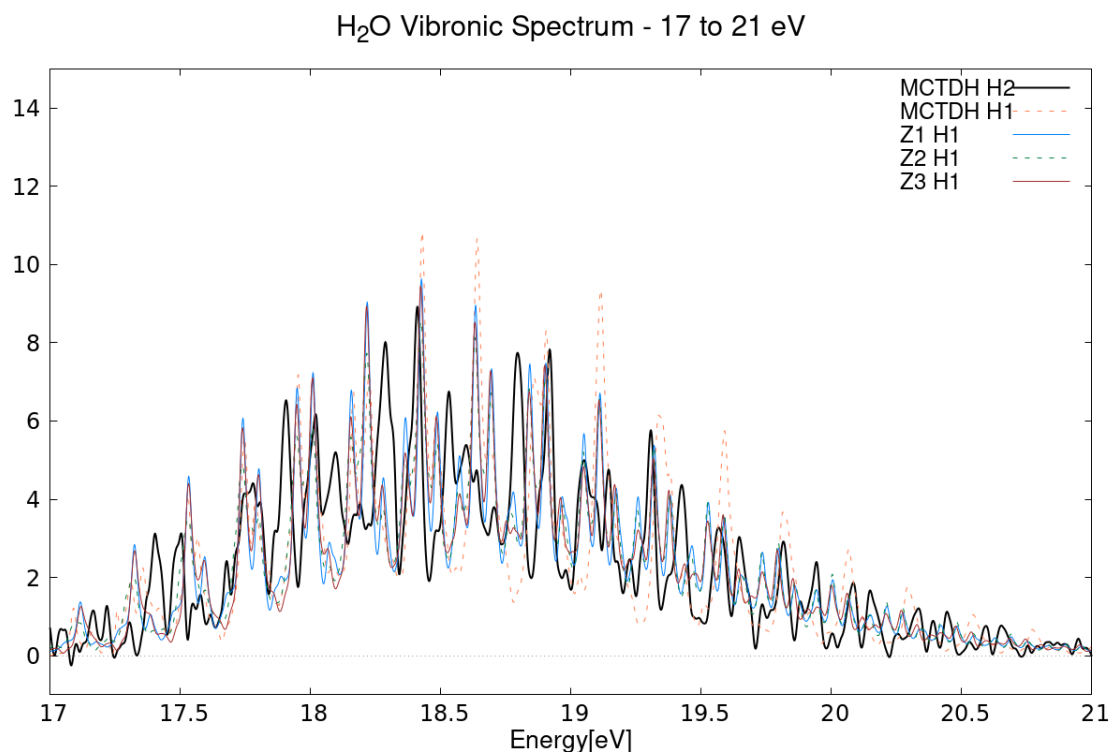


Figure 4.3: Theoretical vibrationally-resolved photoelectron spectrum of H_2O between 17–21 eV. A higher resolution photoelectron spectra with zoomed-in regions is available in Ref [79].

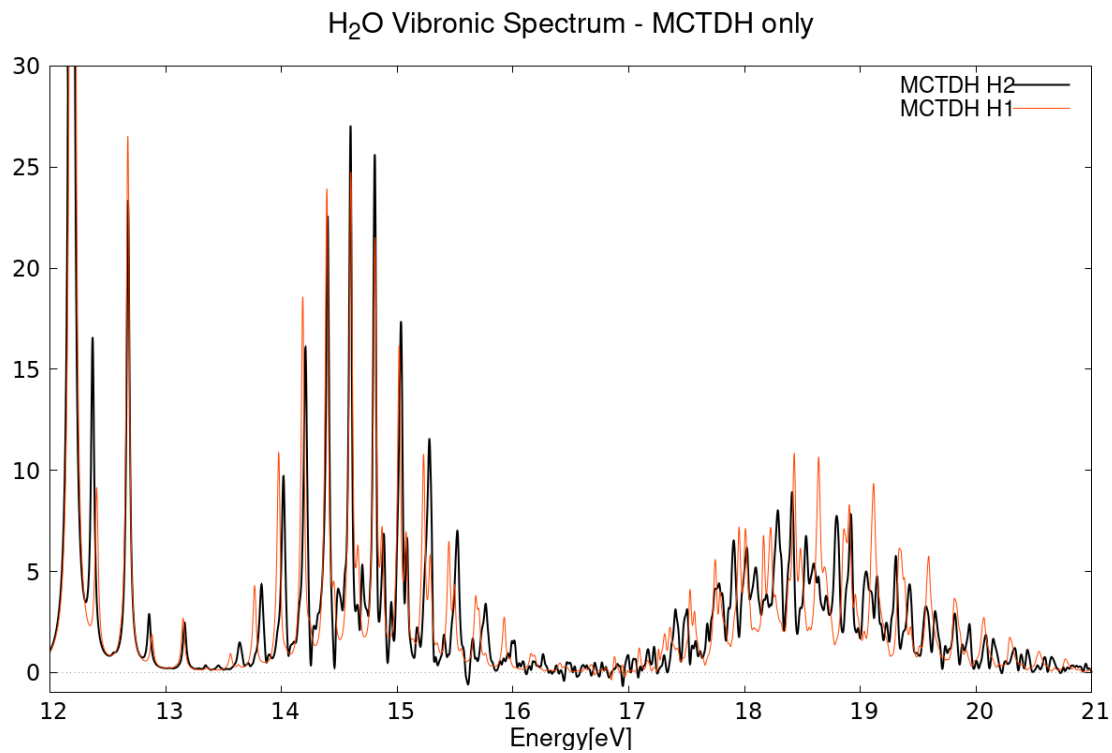


Figure 4.4: Theoretical vibrationally-resolved photoelectron spectrum of H_2O using MCTDH.

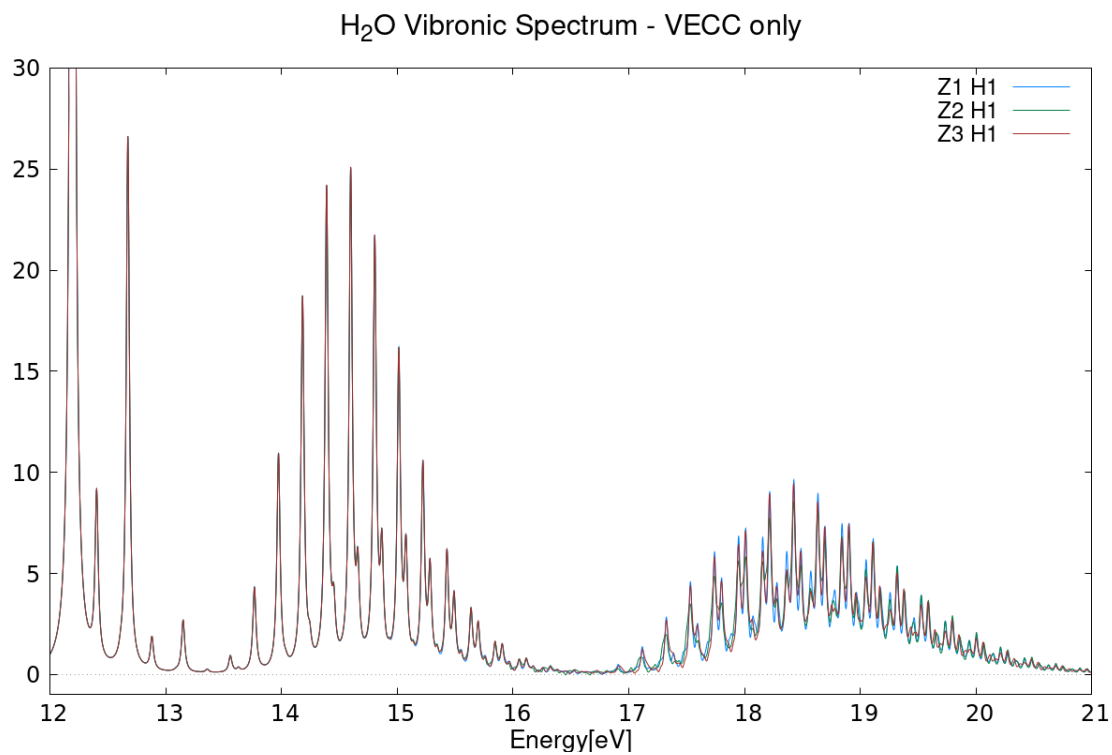


Figure 4.5: Theoretical vibrationally-resolved photoelectron spectrum of H_2O using VECC.

4.2.2 Ammonia, NH_3

Ammonia was the second test-case vibronic model molecular system. It is actually the first molecule we successfully diabatized with `dist`. We have performed calculations on two models differing by symmetry: D_{3h} and C_{3v} . The D_{3h} version explicitly avoids the troublesome A_1 umbrella mode with out-of-plane distortion by the hydrogens moving in and out of the N-H plane. Although the D_{3h} version was useful for PES and seam space studies [80], the less symmetric C_{3v} is more physically realistic and therefore used for comparison to experimental photoelectron spectra in Refs. [81, 82]. Our vibronic model of NH_3 has 3 electronic surfaces and 6 normal modes, with a full valence active space of 7o7e.

Both electronic transition bands exhibited in the theoretical spectra Figure 4.6, appear to match the experimental spectra position-wise and shape-wise. The broad high energy band centered around 16.3 eV, corresponding to the degenerate ${}^2E$ state, is a bit lower in intensity than expected⁵. The $\tilde{B}^2A_1$ hump with its shoulder starting at around 25 eV in Edvardsson et al.’s study [82] was not covered in our simulations, as it is too high in energy.

A major difference in the VECC spectra is the stark lack of vibronic fine structure in this high energy band, as all Z1 to Z3 spectra are relatively smooth. Whereas, all simulations have reproduced the lower energy band fairly accurately. This is a known issue for VECC [9], and we are iterating on methods to resolve the peak to match MCTDH more closely. The quadratic (H2) MCTDH simulation has comparatively smaller amplitude splittings when compared to the linear order spectra. VECC was unable to run quadratic NH_3 models due to convergence issues.

⁵Though it can be tuned by playing around with τ parameter in `autospec`.

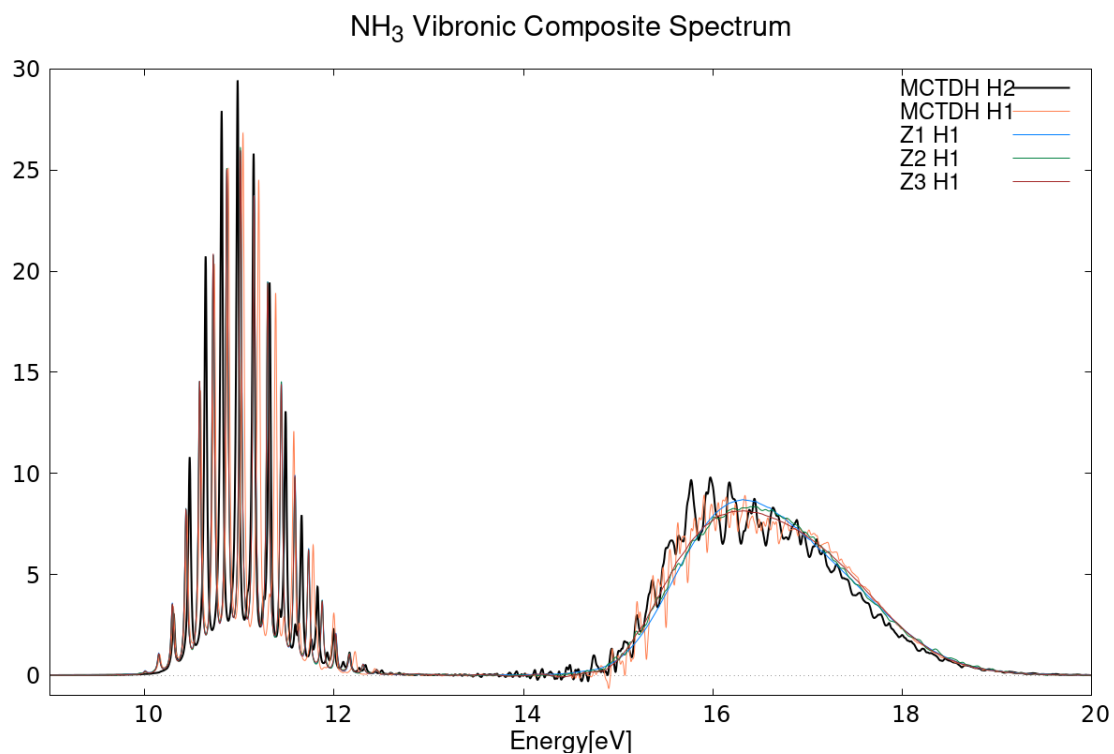


Figure 4.6: Theoretical vibrationaly-resolved photoelectron spectrum of NH₃.
NH₃ Vibronic Spectrum Comparison - 14 to 19 eV

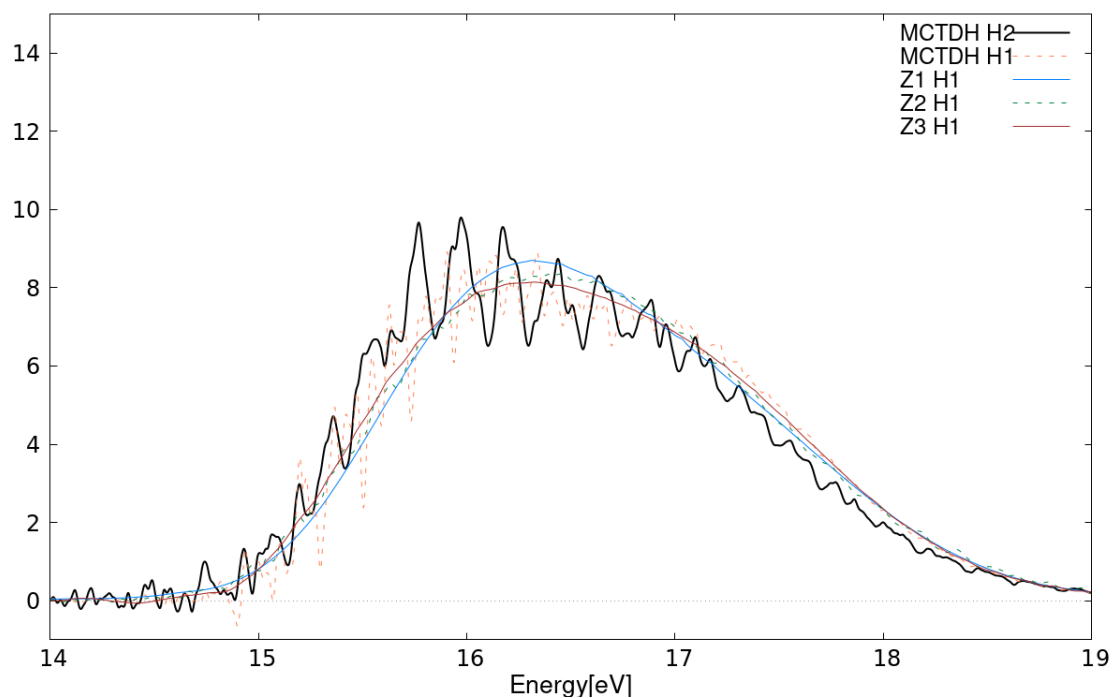


Figure 4.7: Theoretical vibrationaly-resolved photoelectron spectrum of NH₃ between 14–19 eV.

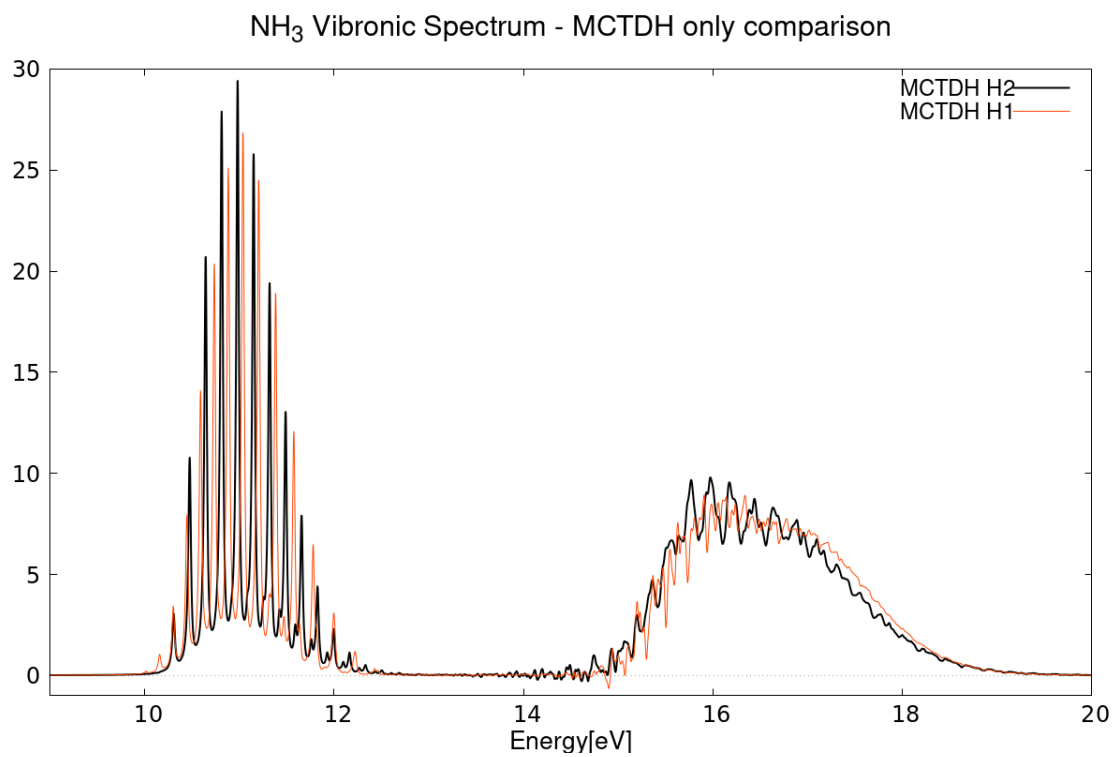


Figure 4.8: Theoretical vibrationally-resolved photoelectron spectrum of NH₃ using MCTDH.

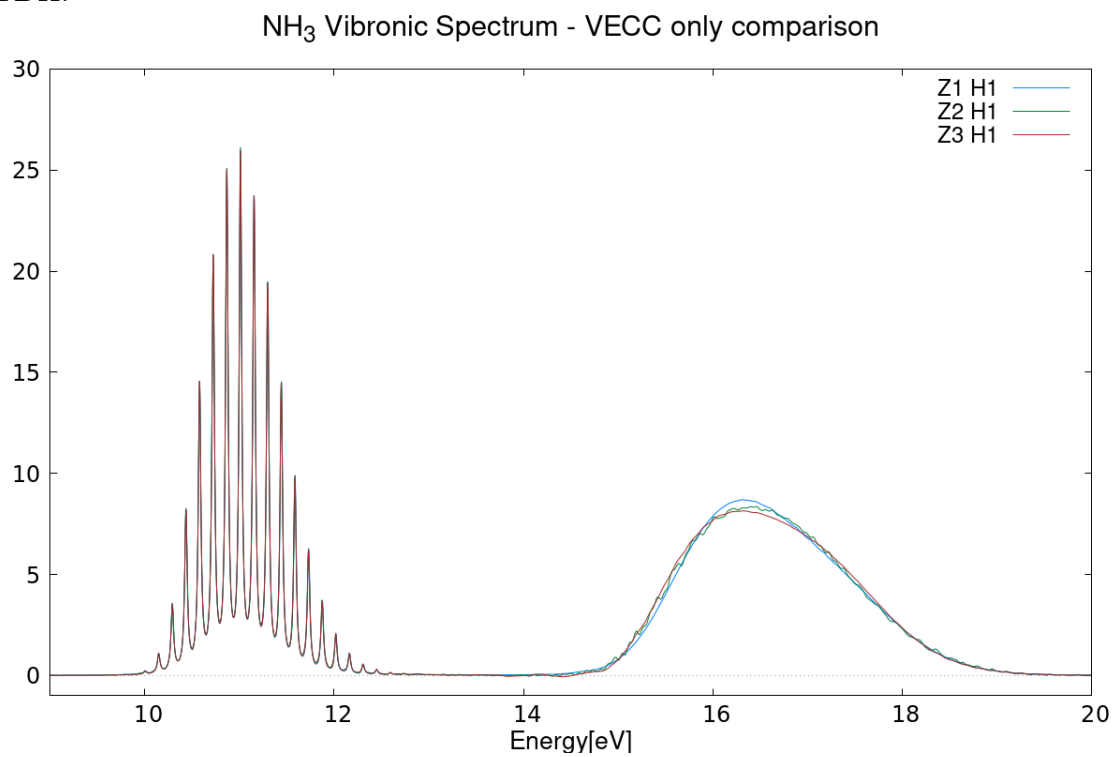


Figure 4.9: Theoretical vibrationally-resolved photoelectron spectrum of NH₃ using VECC.

4.2.3 Phosphine, PH_3

Phosphine was the third test case molecule. Featuring similar (C_{3v}) symmetry and being a pnictogen trihydride compound one row below ammonia, this was deemed to be another compatible test subject to apply our code towards. Predictably, the phosphine theoretical photoelectron spectra shares the same ammonia-like double hump profile. However, as compared to the experimental by Maripuu et al. in Ref. [83], we had to red-shift our spectra by approximately 3 eV for alignment. The cause behind this deviation is unknown, but likely due to mistakes in selecting appropriate states⁶. The exact same active space as ammonia (7o7e) was used. VECC's Z3 simulation in Figure 4.14's pink has ostensibly managed to catch some ripples within the far UV vicinity. The interpretation here supports that as you increase the accuracy of the calculation in VECC, the more it approaches MCTDH's accuracy. Once again, PH_3 just like the previous two examples is a relatively small system, hence we are able to propagate for very long 1000 fs runs. In reality, this is excessive. Most quantum dynamical studies on phosphine cation, such as by Bhattacharyya, are nominally only around 200 fs in length [84].

⁶With our implementation of `dist`, it is easy for the user to set a predefined linear shift to correct this issue (if desired)

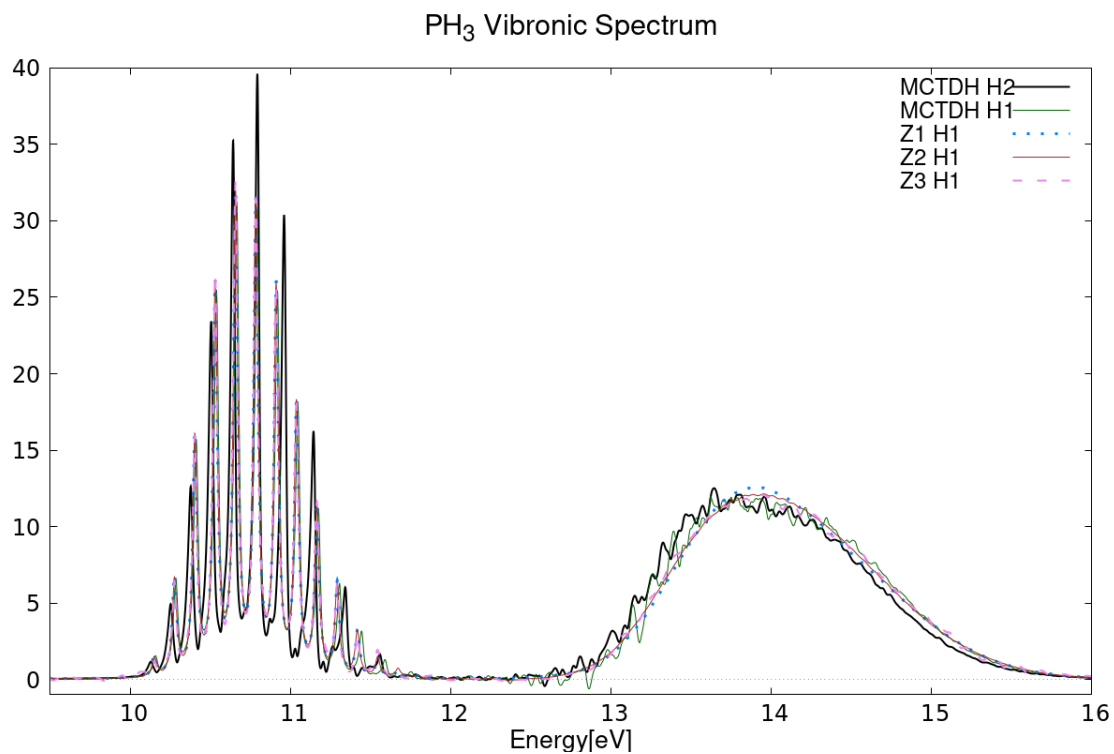


Figure 4.10: Theoretical vibrationally-resolved photoelectron spectrum of PH₃.
PH₃ Vibronic Spectrum - High Energy Peak

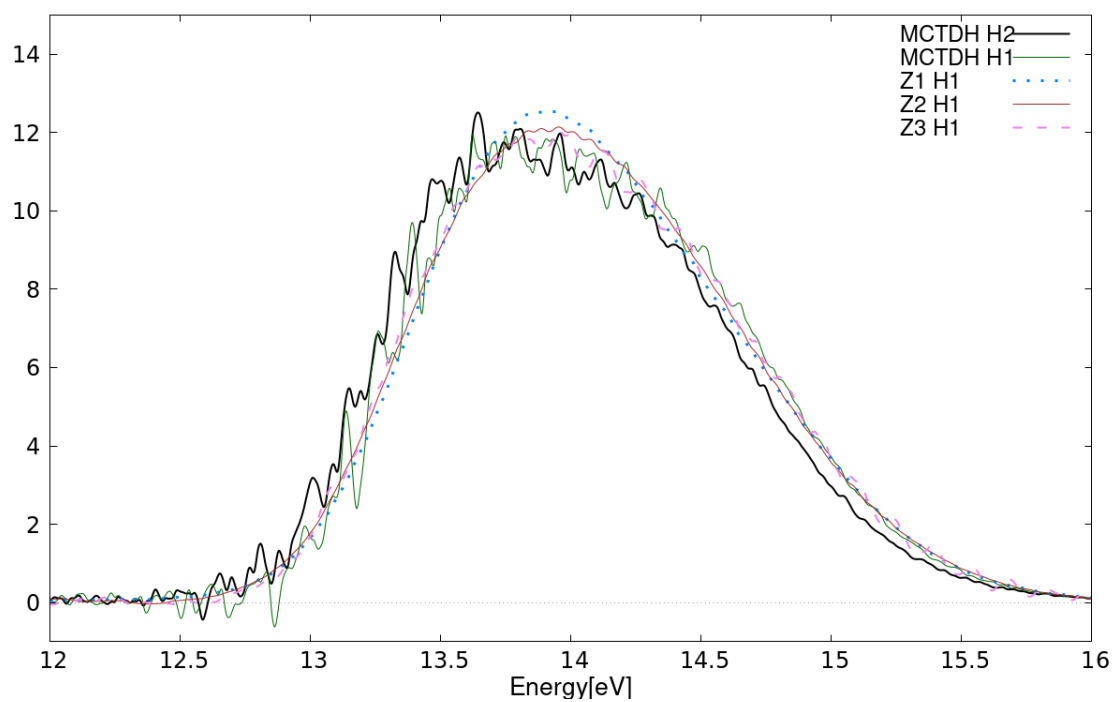


Figure 4.11: Theoretical vibrationally-resolved photoelectron spectrum of PH₃.

4.3 Complex Systems

In Section 4.2, we demonstrated the reliability of the protocol in capturing the qualitative shape of the bands. Following this, we moved onto transition metal complex systems. Titanium tetrachloride and caffeine (from Section 1.3) were not re-tested with the new Zeng vibronic models, due to aiming for newer systems that we have never simulated before. As previously described in Section 1.2, open-shell transition metal complexes are of high importance because they include many interesting and rich chemical properties [52, 85]. We began by investigating CoF_3 and RhF_3 , as these transition metal tri-fluorides (MF_3) are well known to show excellent vibronic, JT, and SOC effects as cited by Mondal. They are also comparatively more accessible to simulate. We then examined $\text{Fe}(\text{CO})_5$, which was of interest due to its unstable nature, and difficulties other researchers had in simulating its spectra [41, 86, 87, 88].

4.3.1 Cobalt trifluoride, CoF_3

CoF_3 was the first complex transition metal system we used to test **dist**. This SOC model has 15 electronic surfaces⁷ and 6 normal modes. Our active space of CoF_3 is a reduced (from full-electron and full valence) 5o6e. Since this was the first system with a metallic atom, we depart from using the standard **cc-pVDZ/cc-pVTZ** basis sets and instead use the Sapporo family of relativistic bases **SPKrnZP** ($n = \text{T}$) [89]. A key observation was that the active space of the cobalt system was difficult to converge when compared to the simple test case molecules. Therefore, we employed a strategy where we slowly filled bunches (3–5) of electrons into orbitals and slowly built up the complete active space. After every optimization run, we needed to preserve the orbital ranking, this meant turning on the **norder** switch in GAMESS and playing around with **iorder** to reorder orbitals. Primarily, this effort was done to ensure that the valence d-orbitals were always the HOMOs.

⁷ $(^5A'_1$ ground state [9] with occupation scheme of $[\text{core}](a'_2)^2(e'')^2(e')^2$)

Finding experimental spectra for CoF_3 was difficult, and we eventually ended up using Mondal et al.'s 17o30e 2 states and 3 modes theoretical spectra [39]. Note that this spectra is of a very reduced system (2,3) compared to our full vibronic model (15,6) and so we do not expect to match quantitatively. The Chebyshev propagation method used in the original study was also not used in our approach [3, 39]. Rather, the reason for comparison is to highlight the contributions of SOC and active JT effects.

The simulated spectra in Figure 4.12 has photoelectron bands centered at 0.16 and 0.67 eV. All propagations agree on the leftmost peak, whereas the bands seen in the 0.67 eV range show meaningful variations. We have introduced a linear shift of 1.21 eV so that the spectrum can be in the positive energy vicinity. Although our spectra may once again be red-shifted, there is general agreement in the presence of two intense bands when comparing to Mondal and collaborators' spectra (see Ref. [39]). Comparing the SOC to non-SOC lines indicates that the SOC introduces a broadening of the lines. This is clearly observed in the constants (H0) spectra. The simulated broadening effect is consistent with the reference spectra. Upon inclusion of higher order expansions (H1 and H2), the 0.67 eV band appears to have split due to JT activity, giving rise to a major peak at 0.60 eV and a minor shoulder starting around 0.75 eV. With Figure 4.18 and 4.19, the origin of the new noticeable hump appearing at 0.75 eV attributed to SOC is more flattened and plateau-like in VECC. The VECC spectra are also noted to be more resolved in the 0.75 eV region, especially the linear non-SOC result. Relative intensity of MCTDH propagations were also consistently lower than VECC within the higher energy range.

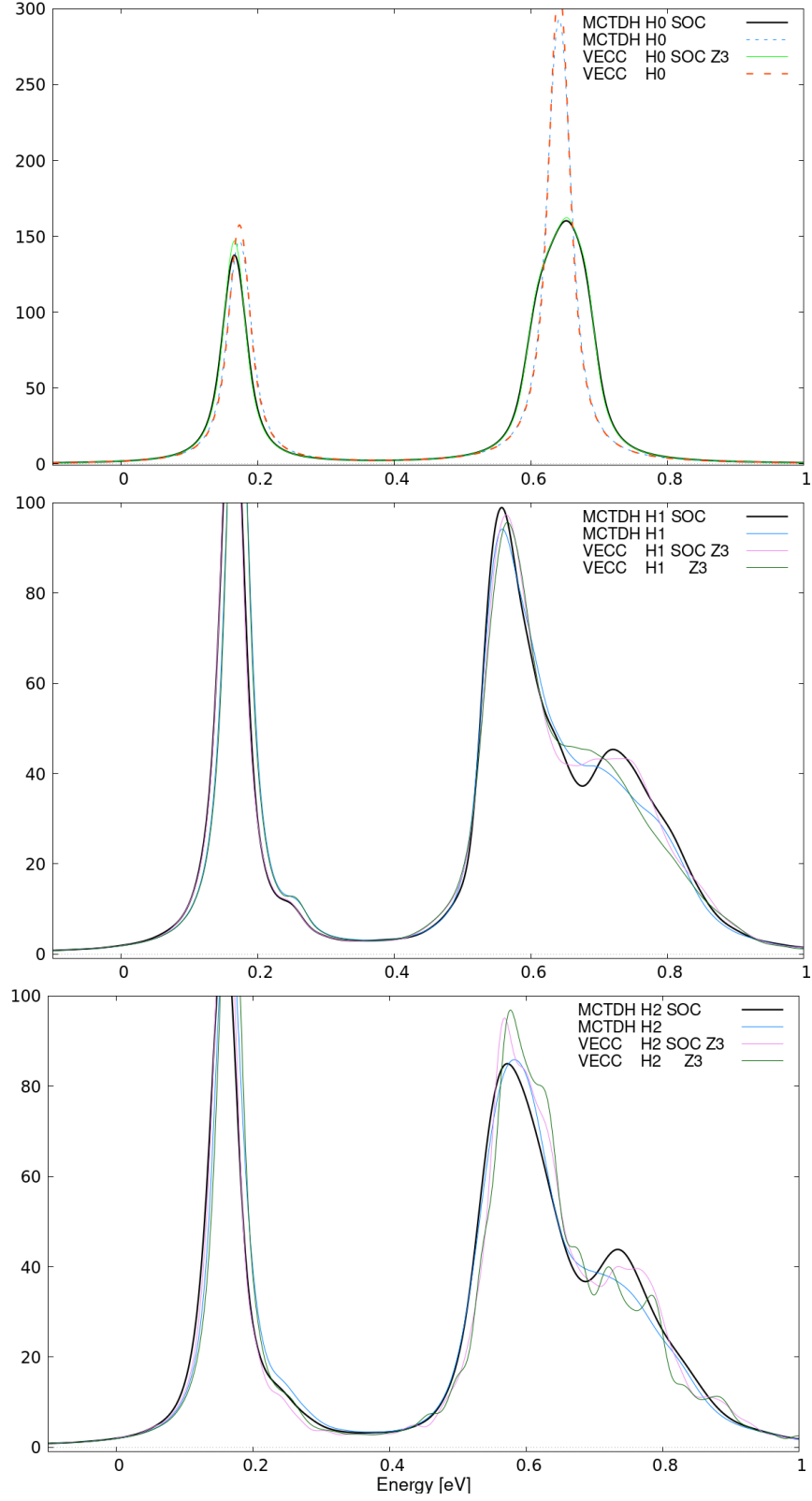


Figure 4.12: Theoretical vibrationally-resolved spectra of CoF_3 for three truncations of the vibronic model expansion. The upper panel includes only constant order contributions, the middle panel up to linear order, and the bottom panel up to quadratic order.

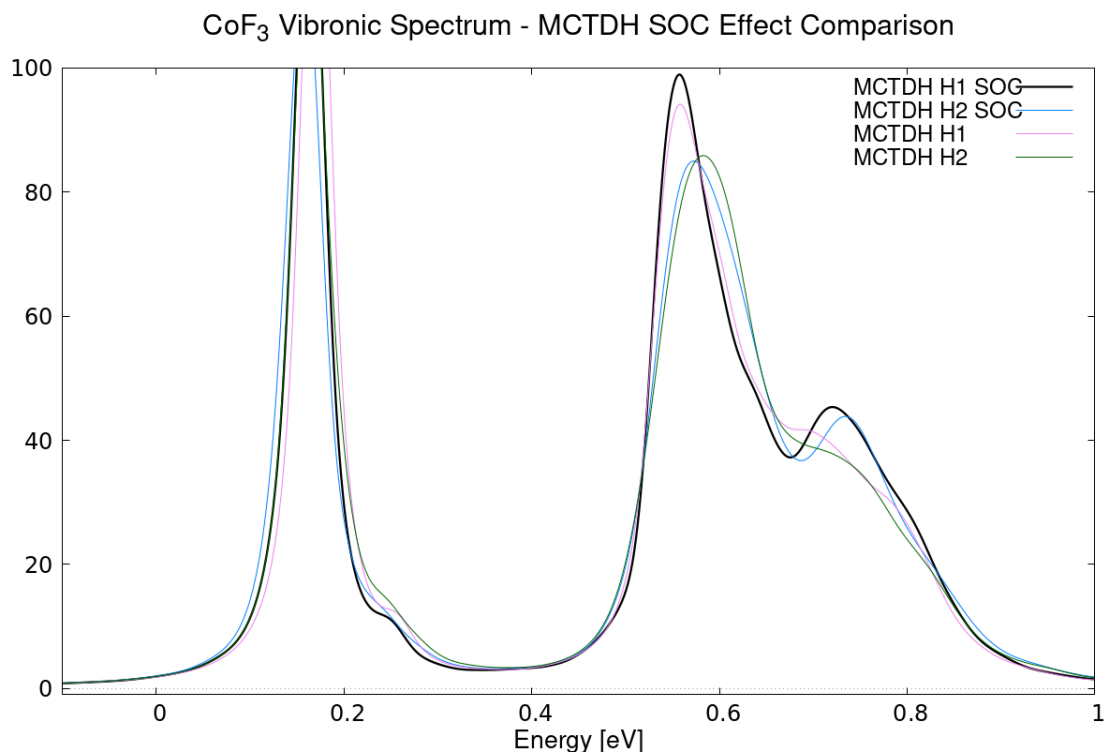


Figure 4.13: Theoretical vibrationally-resolved spectrum of CoF₃ using MCTDH.

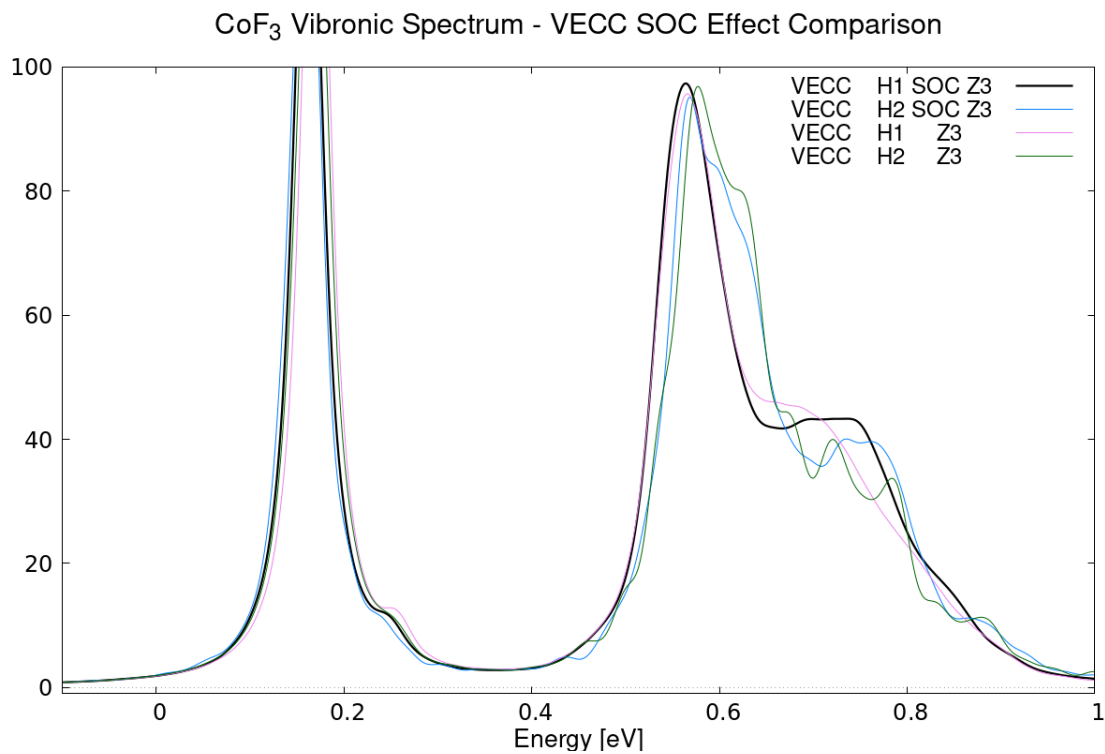


Figure 4.14: Theoretical vibrationally-resolved spectrum of CoF₃ using VECC.

4.3.2 Rhodium trifluoride, RhF_3

RhF_3 is another transition metal trifluoride, and is expected to experience a more prevalent SOC effect. Expectedly, the SOC terms in the model are on average one magnitude higher (in eV) as compared to CoF_3 . The active space of this model is a repeat of CoF_3 , 5o6e. Likewise, this model also has 15 electronic states and 6 normal modes and required a linear shift of 1.22 eV. In parallel to the earlier pnictogen hydrides, CoF_3 and RhF_3 both share a similar absorption band profile shape and width.

Simulated spectra at three different truncation's of the Hamiltonian with and without SOC corrections are shown in Figure 4.15. In the upper panel (constants only H0) the SOC corrections produce significant changes in the band structure. The single peak, with SOC, dramatically splits into four peaks within the 0.55 eV range. Both MCTDH and VECC (Figures 4.16 and 4.17) have excellent agreement here. This pronounced splitting has evidently broadened the higher energy peak into two discernible humps of equivalent intensity, upon inclusion of higher order terms. It has also lowered the intensity around the 0.60 eV range. Without SOC, the higher energy band remains similar in shape to CoF_3 . It is difficult to assign whether the Figure 4.17 VECC calculated spectra features emergence of another true new peak at 0.70 eV. The low energy peak centered around 0.10 eV also has a telltale red-shift due to SOC, it also slightly broadened and decreased in intensity.

In contrast to CoF_3 , comparison to experimental or literature-simulated spectra will not be made, as records are scarce⁸. Besides, these comparisons would be misleading as our chosen D_{3h} geometry is somewhat artificial. In reality the crystal structure of MF_3 compounds more likely takes on octahedral geometries [90, 91]. RhF_3 has also confirmed the repeated success of adding SOC into our models. As referenced in Chapter 1, we were not capable of achieving transition metal complex simulations when applying Santoro models.

⁸Only experimental spectra available were diffuse reflectance and X-ray K-edge absorption spectra by [90] and [91] respectively.

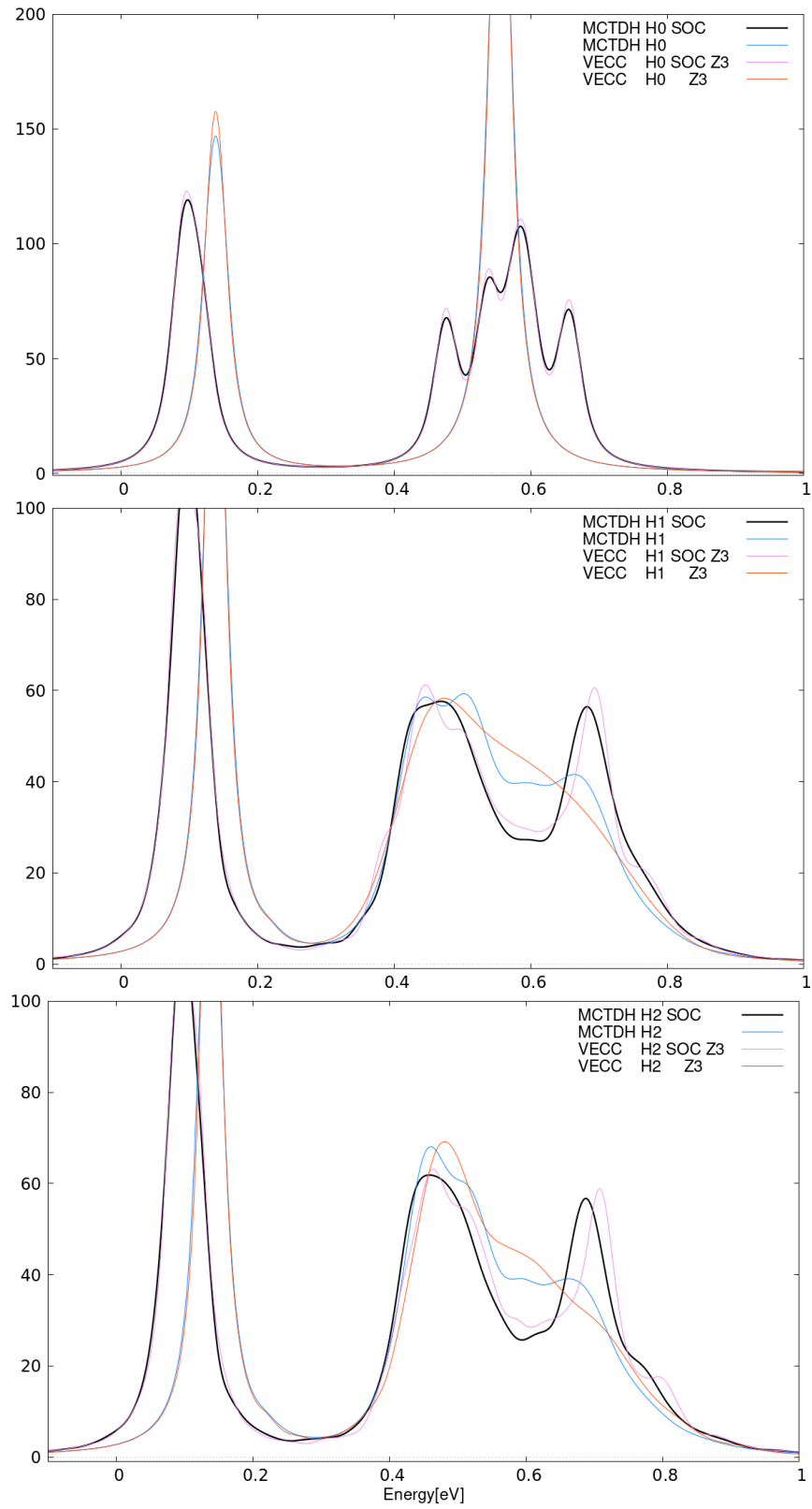


Figure 4.15: Theoretical vibrationally-resolved spectra of RhF_3 for three truncations of the Hamiltonian expansion. The upper panel includes only constant order contributions, the middle panel up to linear order, and the bottom panel up to quadratic order.

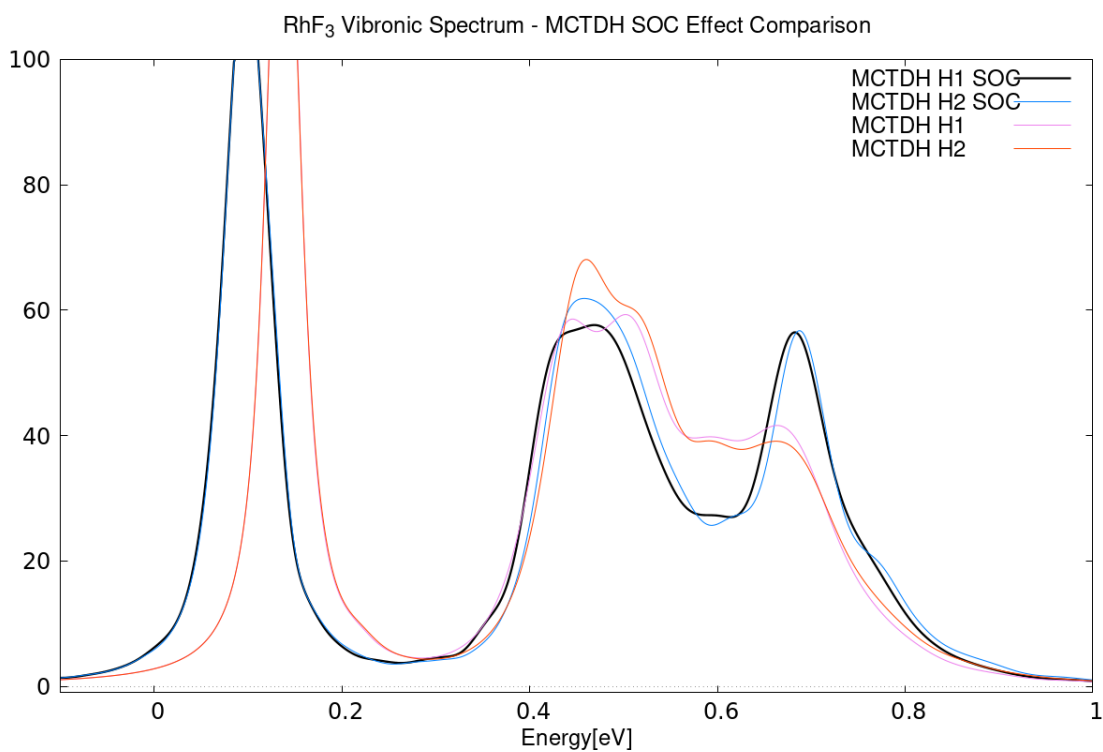


Figure 4.16: Theoretical vibrationally-resolved spectrum of RhF_3 using MCTDH.

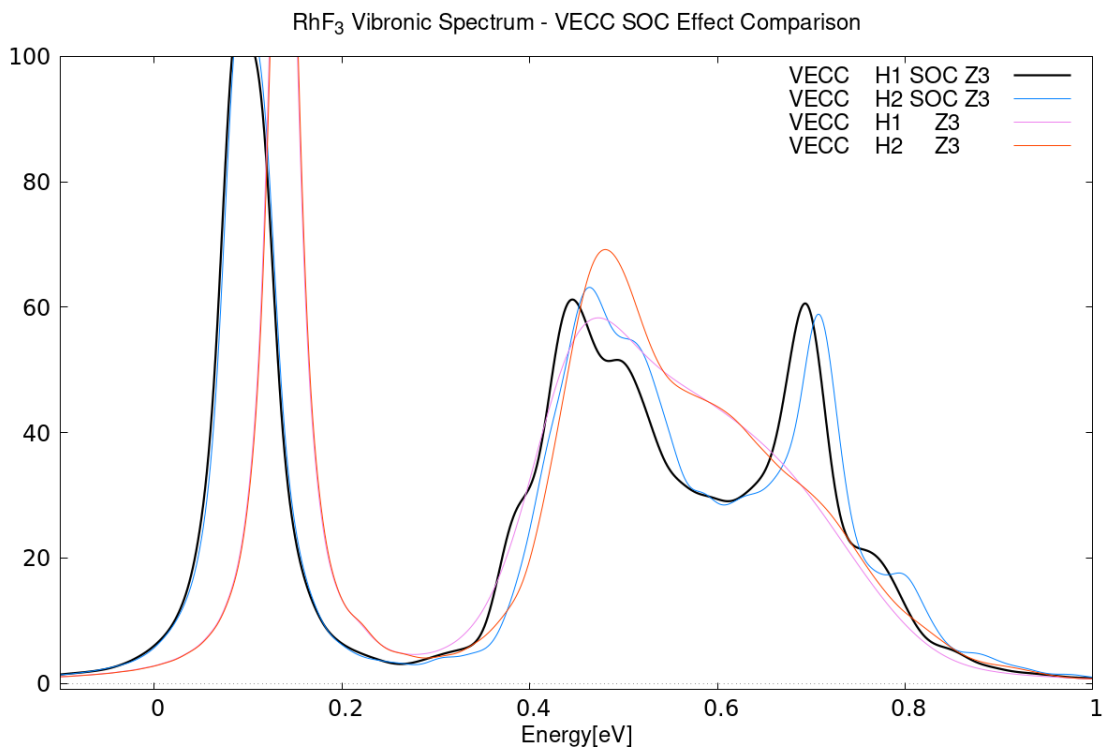


Figure 4.17: Theoretical vibrationally-resolved spectrum of RhF_3 using VECC.

4.3.3 Iron pentacarbonyl, $\text{Fe}(\text{CO})_5$

Iron pentacarbonyl is the largest molecular system we have studied. With $\text{Fe}(\text{CO})_5$'s high symmetry and nearly-degenerate states, the prospects for exhibiting properties of chemical interest in particular for photochemistry applications is well-known [41, 86, 87]. This model has 17 electronic surfaces (12 triplet and 5 singlet states) and 27 normal modes. There are numerous near-degenerate states, due to $^3E''$, $^3E'$, and $^1E''$ all being close in energy, it was expected that there will be strong SOC between $^3E'$ and $^1E''$. We used an active space of 8o10e for $^1A'$ $\text{Fe}(\text{CO})_5$.

We were unsuccessful in reaching the all-electron full valence active space (recognized as 11o18e) due to $\text{Fe}(\text{CO})_5$'s notable instability. Regarding the active space problem, just like CoF_3 , we could not do a one-shot calculation. Instead, we decided to use the bunching strategy of slowly filling up the orbitals with bunches of electrons and orbitals at a time. The initial calculations were cationic and as we approached the neutral molecule, a roadblock appeared between a set of a'' and e' of orbitals 46 through 48. Eventually further calculations resolved the orbital ranking: a'' was lower than e' . Finally, the fully electron-filled neutral molecule yielded the an equilibrium geometry with an active space of 8o10e as illustrated in Figure 4.18. Expanding to a larger active space proved to be trouble-making, as the instability itself annihilates the convergence in geometry optimizations. Correspondingly, Nakano et al. has noted this same instability issue in their FeCO^+ wavefunction study [38]. Therefore, this has stymied our many efforts and calculations beyond 8o10e.

After figuring out the active space and completing a template, another major problem confronted involved the size of the model. It simply had too many normal (27) modes⁹. This meant that the MCTDH84 calculations would frequently fail due to inability in handling the massive amount of vibronic coupling terms given. Our solution was to implement in

⁹Typically, MCTDH studies often employ the multi-layer MCTDH approach to tackle such systems with large degrees of freedom [19]

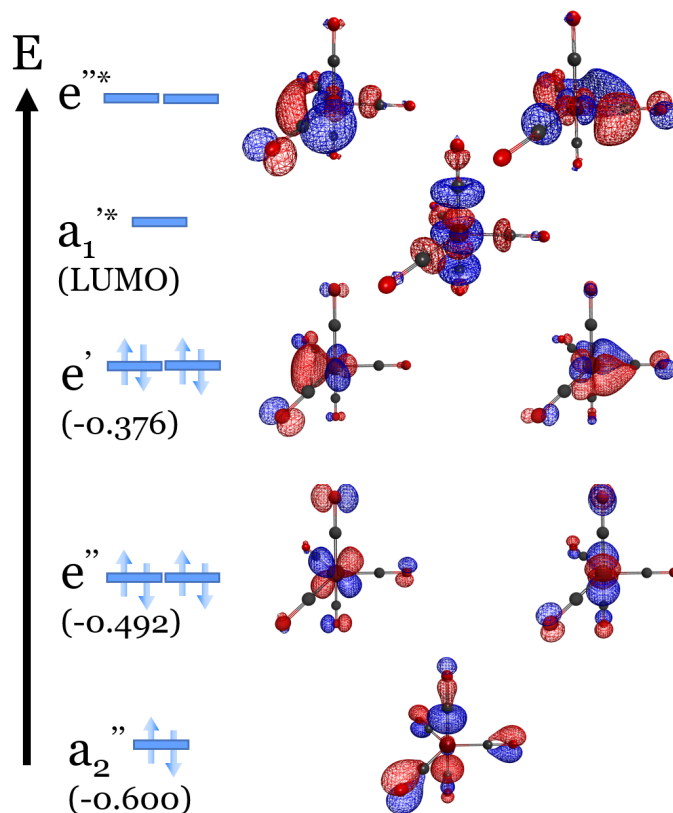


Figure 4.18: The ${}^1A'$ active space of 8010e $\text{Fe}(\text{CO})_5$.

dist a screening subroutine (Section 3.3.3). This allowed us to filter¹⁰ the most relevant 12 modes out of 27, and build a truncated Hamiltonian that in theory retained its most important couplings. For $\text{Fe}(\text{CO})_5$, this delivered a highly useful determination of the most important modes ([7, 8, 12, 13, 15, 18, 19, 21, 22, 23, 30, 31]) for inclusion out of the full 27. The selected twelve modes and their vectors are depicted in Figure 4.19. In the end, after contacting Prof. Dieter Meyer, he revealed to us that our workaround was in actuality not needed, as MCTDH86 was the answer to sizeable calculations (Section 3.3.4).

The simulated $\text{Fe}(\text{CO})_5$ spectra profile broadly features a highly resolved band around 3 eV and a double hump absorption band within the 6-12 eV range as shown in Figure 4.23 and Figure 4.24. The SOC effect is only apparent in the constant order spectra, Figure 4.21 and Figure 4.22. Unlike the clear SOC effect occurring in the two trifluoride examples,

¹⁰based on a predefined threshold value of -0.60 in magnitude

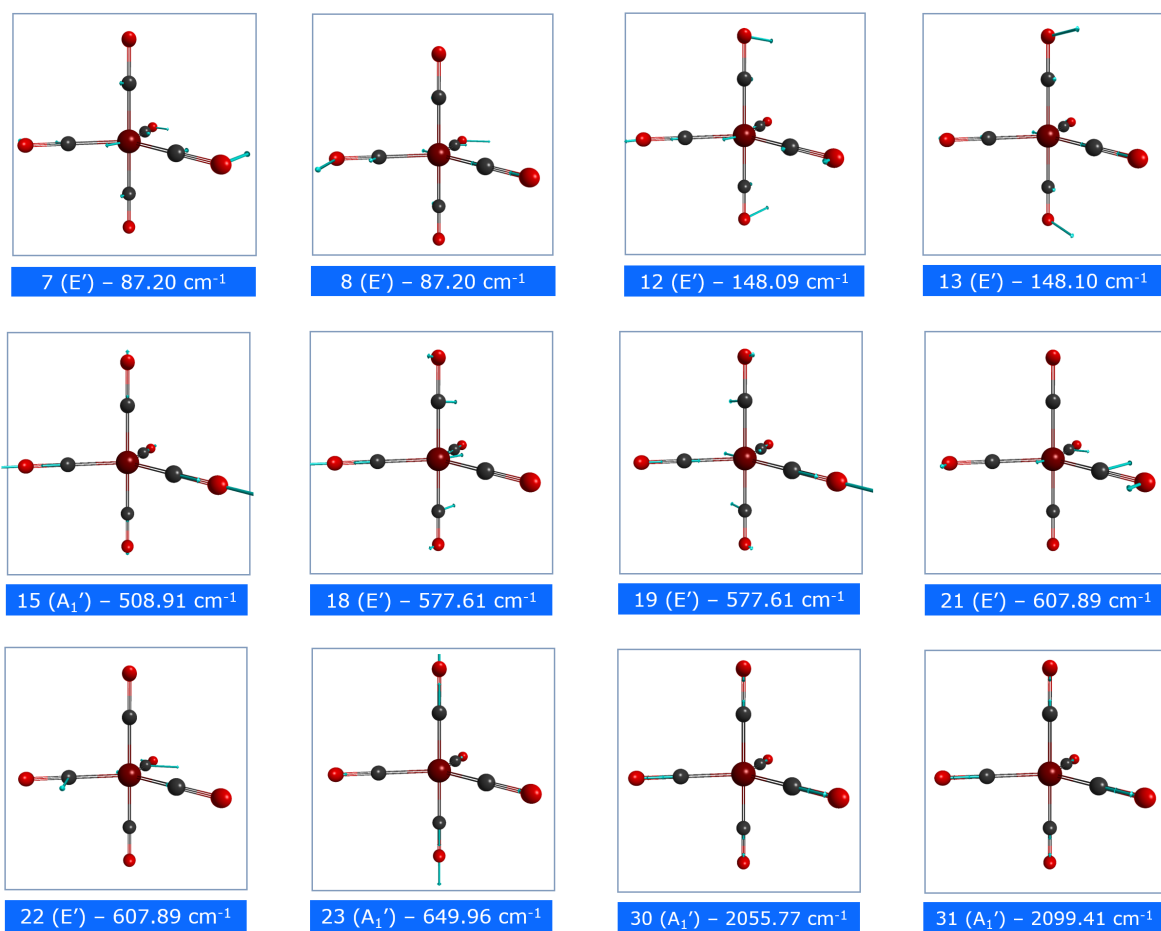


Figure 4.19: 3D view of the chosen 12 screened-modes (out of 27 possible modes) of $\text{Fe}(\text{CO})_5$ and their associated frequencies. The modes were filtered by means of an effective vibronic coupling magnitude calculation.

despite having SOC matrix elements terms comparable in magnitude (10^2 cm^{-1}) as CoF_3 , $\text{Fe}(\text{CO})_5$'s SOC ultimately only contributes to about 150 cm^{-1} or 0.02 eV in splitting. This was determined after the fact by computing the SOC eigenvalue spectrum in Figure 4.20. In $\text{Fe}(\text{CO})_5$ simulated photoelectron spectra studies by Gaba et al. [86], the computed SOC splitting can contribute as high as $\sim 1.2 \text{ eV}$. However it is noted that their study is regarding photodissociation and that the active space and/or level of theory are different. The SOC in constants (H0) spectrum reduced the intensity of peaks at 8.5 and 9.8 eV, but does not affect peaks at 9.4 or 10.7 eV. By comparison to the trifluorides, the SOC broadening is marginal—the shape remains almost exactly the same, with SOC just slightly outlining the non-SOC. The SOC effect in linear and quadratic order spectrum are undetectable, and is likely due to being quenched by inclusion of higher order terms and JT [40]. Contributions from higher order SOC terms and their effect on spectra was always less compared to constant order SOC. Therefore, we can conclude that constant order spectra are critical in indicating the predicted extent of SOC effect (in higher order spectra), and that SOC can be ruled out in playing an extensive role in our $\text{Fe}(\text{CO})_5$ model. Certainly, the SOC of the fourth row element Fe is not strong. It is only enough to slightly split each of the two ^3E -type terms to three levels. Hypothetically, JT activity appears to dictate the primary interest instead.

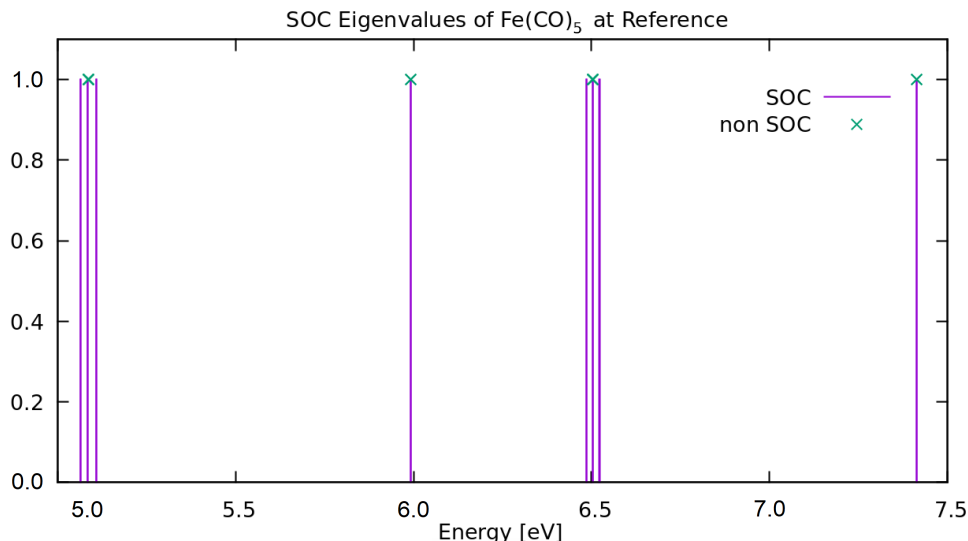


Figure 4.20: SOC eigenvalues spectrum for $\text{Fe}(\text{CO})_5$ *refG* GMC-QDPT energies.

The H1 linear order MCTDH spectra in Figure 4.23 spectra has more vibronic fine structure as compared to the H2 quadratic order spectra in Figure 4.24. It is speculated that the reduced degrees of freedom in this large model resulted in less smoothness in the peaks, allowing the linear vibronic fine structure to be resolved. Inclusion of quadratic (H2) order terms interestingly quenches the resolved fine structure in all cases. VECC unfortunately did not reproduce any fine structure within the same double hump vicinity of 6-12 eV. This has similarities with the behaviour of VECC exhibited in earlier spectroscopies of NH_3 , where the lower energy band was better resolved. Many VECC propagations were also observed to have difficulties in convergence for $\text{Fe}(\text{CO})_5$, especially in the quadratic case, which can explain the approximate 1 eV red-shifted discrepancy in the higher energy band seen in Figure 4.24.

Comparison to experimental or other simulated spectra will not be made extensively. Justification in the general band shape of our simulated spectra can be argued by comparison to the lower energy 8-11 eV flank of $\text{Fe}(\text{CO})_5$ experimental spectra shown in Fig. 5 of Ref. [41]. However, this agreement might be coincidental, as the study is concerned with photodissociation, which we have not incorporated. We anticipated the inclusion of SOC to noticeably modify the spectrum. However, as demonstrated above in the H0 spectrum, the SOC of *Fe* was not significant enough to meet our anticipation.

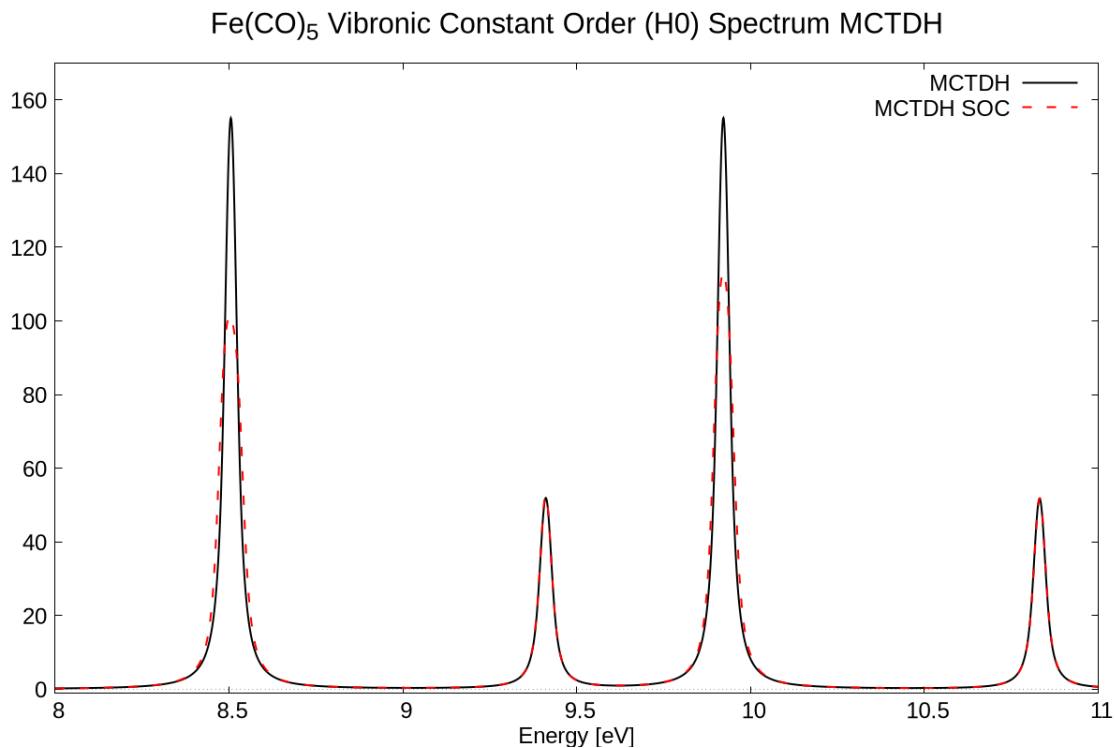


Figure 4.21: Theoretical vibrationally-resolved spectrum of Fe(CO)₅. Lower energy peak not shown. MCTDH comparison of SOC effect. Propagated for 180 fs and resolved with $\tau = 40$.

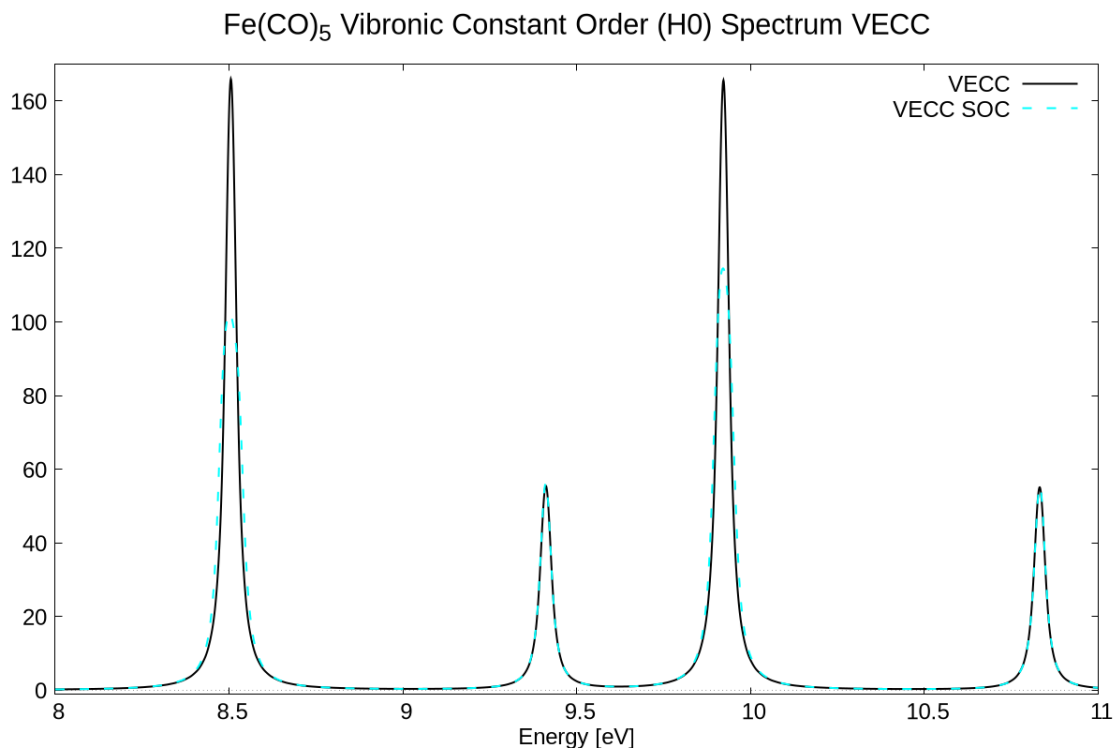


Figure 4.22: Theoretical vibrationally-resolved spectrum of Fe(CO)₅. Lower energy peak not shown. VECC comparison of SOC effect. Propagated for 180 fs and resolved with $\tau = 40$.

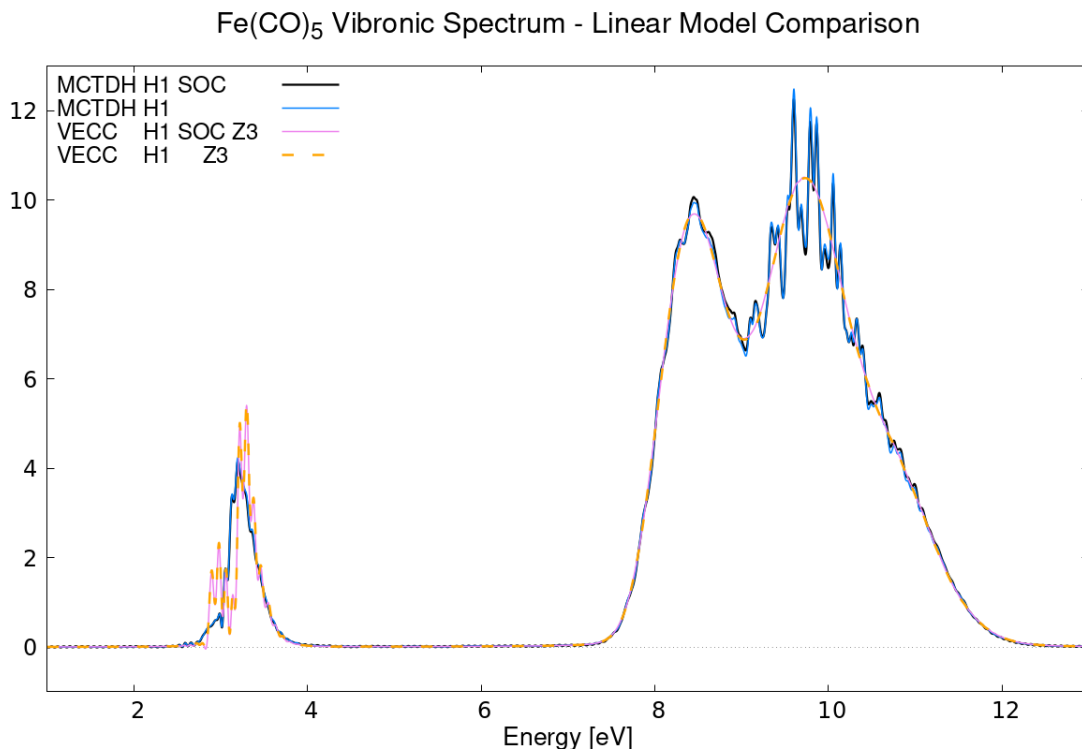


Figure 4.23: Theoretical vibrationally-resolved linear order spectrum of Fe(CO)₅. Propagated for 180 fs and resolved with $\tau = 40$.

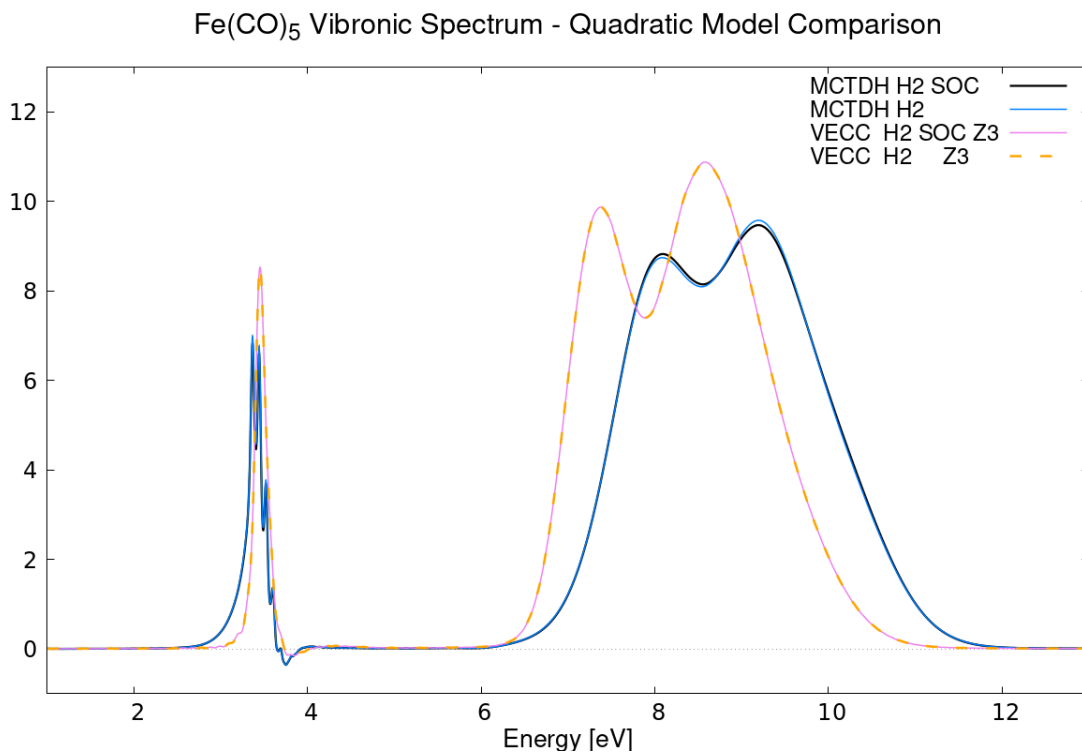


Figure 4.24: Theoretical vibrationally-resolved quadratic order spectrum of Fe(CO)₅. Propagated for 180 fs and resolved with $\tau = 40$.

Chapter 5

Conclusion

A new vibronic spin-orbit coupling model protocol suitable for spectroscopy was presented and discussed. An overview of vibronics, the current state of vibronic model research, along with a concise explanation to the thesis project objectives were introduced in Chapter 1. The mathematization and methods used to construct our vibronic models were provided in the content of Chapter 2. Development of the diabaticization code was discussed in Chapter 3, and excerpts from a $\text{Fe}(\text{CO})_5$ calculation were used to highlight the connections between theory and individual calculations, as well situate them in the larger picture. Results for six systems were presented in Chapter 4. Basic test-case models were used to establish baseline parameters and verify the validity of the approach in concurrent simulations between MCTDH and VECC. Larger D_{3h} systems such as CoF_3 , RhF_3 , and $\text{Fe}(\text{CO})_5$ were used to show the practicality and accuracy of the method in capturing both SOC and JT effects.

The protocol and newly developed diabaticization code has extended the previous Zeng spin-orbit diabaticization protocol to integrate spectroscopic applications. The Python software package, `VEModel`, which includes the main diabaticization program `dist` and utility program `DiabToolkit`, produced in this work automated the process to increase efficiency and accessibility. With this thesis work, we have connected a clear, streamlined path from input geometry to vibronic model Hamiltonian construction in GAMESS, quantum dynam-

ics simulation by choice of MCTDH or VECC approach, and ultimately generation of the desired vibronic absorption spectra, even for transition metal complexes. We also extended the VECC method to support SOC corrections. Extraction and inclusion of diabatic SOC matrix elements has led to the improvement of the VECC approach to accept complex-valued matrices as input.

Results of using the diabatization code has led to corroboration of generalized pnictogen hydride and CoF_3 spectra of previous studies. The lack of vibronic fine structure in our spectra can be attributed to using full normal modes, which often makes spectra very smooth. Our protocol implementation enables (up to) full modes and arbitrary states simulations, which is a considerable advantage. Although the spectra are not in complete agreement, the overall absorption band shape and width are reproduced accurately. SOC was especially prominent and effective in the CoF_3 and RhF_3 cases. It was found that in the final $\text{Fe}(\text{CO})_5$ system, only constant order spectrum showed validation in SOC effect, however minimal at ~ 0.02 eV, and it is presumed that the molecule is more prone to exhibiting Jahn-Teller characteristics. This has also highlighted the importance of H0 constant order vibronic models for indication of yielding SOC effect in non-truncated models.

While our protocol is versatile, it is not perfect, and there are still deficiencies and limitations. For example, the case of cross-correlation functions could signal that there are more oversights. In this project we have recognized that simplifying everything down to a one-size fits all simple model is unrealistic. To use the code requires preparing custom templates, which requires specialized knowledge of GAMESS and active spaces. As one can see in the $\text{Fe}(\text{CO})_5$ example, the full valence active space might not be attainable due to convergence failures. Even though running the code is indeed very accessible, to get through the protocol is challenging. The sheer amount of parameters and caveats each spectrum run demands is not only dizzying, but also means that it is yet another unavoidable issue that requires expertise to satisfy.

Despite these limitations the method overall is extremely successful. Novice (undergradu-

ate level) users gain an accessible tool for creating models, allowing for increased delegation of tasks and concentration on the chemistry. Generation of complicated models (e.g., $\text{Fe}(\text{CO})_5$) is now feasible, and possible within a period of < 24 hours as opposed to a week, as was the speed at the start prior to developing the diabatization code. Finally, by automating many of these tasks, we greatly reduce the chance of user error, saving many days per erroneous calculation. Each series of calculations needed to construct the template as seen in Chapter 3 is numerous and dense. It could have taken for instance five days to get through, and if an error was introduced, it could potentially take an unknown amount of time to resolve. Now that the protocol and code has been developed, these costly mistakes are diminished.

We are confident that our protocol can credibly be used to extrapolate further results. A targeted investigation of applying our models to thermofield-VECC is of interest to resume. The examples of the CoF_3 and RhF_3 spectra makes promising headway into future studies. Moving forward, we aim to simulate a congruent CoF_3 model with two states and three modes, in prospects of more accurately matching Mondal’s spectra. Progress on transition metal complex spectra can also deepen by tackling the TiCl_4 case, as discussed in Section 1.3. Heavier systems, such as $\text{Ru}(\text{CO})_5$, may also be analyzed for evaluation of SOC corrections produced by our protocol. In the end, even though $\text{Fe}(\text{CO})_5$ did not fully satisfy our high expectations of a ‘dream molecule’, the excellent journey in simulating it gave merit to the idea that developing our vibronic spin-orbit coupling protocol was one of the most rewarding projects we could have undertaken.

Bibliography

- [1] Graham A Worth and Lorenz S Cederbaum. “Beyond Born-Oppenheimer: molecular dynamics through a conical intersection”. In: *Annu. Rev. Phys. Chem.* 55.1 (2004), pp. 127–158.
- [2] Chantal Daniel. “Photochemistry and photophysics of transition metal complexes: Quantum chemistry”. In: *Coordination Chemistry Reviews* 282 (2015), pp. 19–32.
- [3] Wolfgang Domcke and David R Yarkony. “Role of conical intersections in molecular spectroscopy and photoinduced chemical dynamics”. In: *Annual review of physical chemistry* 63 (2012), pp. 325–352.
- [4] Thomas J Penfold et al. “Spin-vibronic mechanism for intersystem crossing”. In: *Chemical reviews* 118.15 (2018), pp. 6975–7025.
- [5] David R Yarkony. “Conical intersections: The new conventional wisdom”. In: *The Journal of Physical Chemistry A* 105.26 (2001), pp. 6277–6293.
- [6] H Köuppel, Wolfgang Domcke, and Lorenz S Cederbaum. “Multimode molecular dynamics beyond the Born-Oppenheimer approximation”. In: *Advances in chemical physics* (1984), pp. 59–246.
- [7] Marcel Nooijen and Songhao Bao. “Normal ordered exponential approach to thermal properties and time-correlation functions: general theory and simple examples”. In: *Molecular Physics* 119.21-22 (2021), e1980832.

- [8] Songhao Bao, Neil Raymond, and Marcel Nooijen. “Time dependent vibrational electronic coupled cluster (VECC) theory for non-adiabatic nuclear dynamics”. In: *The Journal of Chemical Physics* 160.9 (Mar. 2024), p. 094105. ISSN: 0021-9606. DOI: [10.1063/5.0190034](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0190034/19705218/094105_1_5.0190034.pdf). eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0190034/19705218/094105_1_5.0190034.pdf. URL: <https://doi.org/10.1063/5.0190034>.
- [9] Songhao Bao et al. “Vibrational Electronic-Thermofield Coupled Cluster (VE-TFCC) Theory for Quantum Simulations of Vibronic Coupling Systems at Thermal Equilibrium”. In: *Journal of Chemical Theory and Computation* 20.14 (2024), pp. 5882–5900.
- [10] LS Cederbaum, H Köppel, and W Domcke. “Multimode vibronic coupling effects in molecules”. In: *International Journal of Quantum Chemistry* 20.S15 (1981), pp. 251–267.
- [11] Marcel Nooijen. “First-principles simulation of the UV absorption spectrum of ketene”. In: *International journal of quantum chemistry* 95.6 (2003), pp. 768–783.
- [12] Johannes Neugebauer, Evert Jan Baerends, and Marcel Nooijen. “Vibronic coupling and double excitations in linear response time-dependent density functional calculations: Dipole-allowed states of N₂”. In: *The Journal of chemical physics* 121.13 (2004), pp. 6155–6166.
- [13] Anirban Hazra, Hannah H Chang, and Marcel Nooijen. “First principles simulation of the UV absorption spectrum of ethylene using the vertical Franck-Condon approach”. In: *The Journal of chemical physics* 121.5 (2004), pp. 2125–2136.
- [14] Anirban Hazra and Marcel Nooijen. “Vibronic coupling in the excited cationic states of ethylene: Simulation of the photoelectron spectrum between 12 and 18eV”. In: *The Journal of chemical physics* 122.20 (2005).
- [15] Anirban Hazra and Marcel Nooijen. “Comparison of various Franck-Condon and vibronic coupling approaches for simulating electronic spectra: The case of the lowest

- photoelectron band of ethylene”. In: *Physical Chemistry Chemical Physics* 7.8 (2005), pp. 1759–1771.
- [16] Julia Endicott. “Generating vibronic coupling models and simulating photoelectron spectra”. B.Sc. Thesis. 2012.
- [17] Nemanja Golac. “Simulating Vibrationally-Resolved Electronic Spectra Using the Franck-Condon Principle and the Vibronic Model”. B.Sc. Thesis. 2019.
- [18] Daniel Aranda and Fabrizio Santoro. “Vibronic spectra of π -conjugated systems with a multitude of coupled states: A protocol based on linear vibronic coupling models and quantum dynamics tested on hexahelicene”. In: *Journal of Chemical Theory and Computation* 17.3 (2021), pp. 1691–1700.
- [19] Martha Yaghoubi Jouybari et al. “Ultrafast Dynamics of the Two Lowest Bright Excited States of Cytosine and 1-Methylcytosine: A Quantum Dynamical Study”. In: *Journal of Chemical Theory and Computation* 16.9 (2020), pp. 5792–5808.
- [20] Benny Chen. “Molecular Vibronic Models: Using TD-DFT Methods in Gaussian with MCTDH”. B.Sc. Thesis. 2021.
- [21] Tao Zeng. “A diabaticization protocol that includes spin-orbit coupling”. In: *The Journal of Chemical Physics* 146.14 (2017), p. 144103.
- [22] Ekadashi Pradhan and Tao Zeng. “The Lack of Triplet Fusion for an Intramolecular Singlet Fission Chromophore: The Expected, the Unexpected, and a Reconciliation”. In: *The Journal of Physical Chemistry Letters* 15.1 (2023), pp. 43–50.
- [23] Ekadashi Pradhan and Tao Zeng. “Triplet separation after the fastest intramolecular singlet fission in the smallest chromophore”. In: *Journal of Chemical Theory and Computation* 19.7 (2023), pp. 2092–2101.
- [24] Ekadashi Pradhan and Tao Zeng. “Design of the smallest intramolecular singlet fission chromophore with the fastest singlet fission”. In: *The Journal of Physical Chemistry Letters* 13.48 (2022), pp. 11076–11085.

- [25] Achini Japahuge and Tao Zeng. “Theoretical studies of singlet fission: Searching for materials and exploring mechanisms”. In: *ChemPlusChem* 83.4 (2018), pp. 146–182.
- [26] Tao Zeng. “Through-linker intramolecular singlet fission: general mechanism and designing small chromophores”. In: *The Journal of Physical Chemistry Letters* 7.21 (2016), pp. 4405–4412.
- [27] Tao Zeng and Prateek Goel. “Design of small intramolecular singlet fission chromophores: an azaborine candidate and general small size effects”. In: *The Journal of Physical Chemistry Letters* 7.7 (2016), pp. 1351–1358.
- [28] Gregory J Atchity and Klaus Ruedenberg. “Orbital transformations and configurational transformations of electronic wavefunctions”. In: *The Journal of chemical physics* 111.7 (1999), pp. 2910–2920.
- [29] Shaohong L. Li et al. “Model space diabaticization for quantum photochemistry”. In: *The Journal of Chemical Physics* 142.6 (Feb. 2015), p. 064106. ISSN: 0021-9606. DOI: [10.1063/1.4907038](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4907038). eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4907038/14031391/064106_1_online.pdf. URL: <https://doi.org/10.1063/1.4907038>.
- [30] Klaus Ruedenberg and Gregory J Atchity. “A quantum chemical determination of diabatic states”. In: *The Journal of chemical physics* 99.5 (1993), pp. 3799–3803.
- [31] Haruyuki Nakano, Ryuma Uchiyama, and Kimihiko Hirao. “Quasi-degenerate perturbation theory with general multiconfiguration self-consistent field reference functions”. In: *Journal of computational chemistry* 23.12 (2002), pp. 1166–1175.
- [32] Makoto Miyajima, Yoshihiro Watanabe, and Haruyuki Nakano. “Relativistic quasidegenerate perturbation theory with four-component general multiconfiguration reference functions”. In: *The Journal of chemical physics* 124.4 (2006).

- [33] Ryo Ebisuzaki, Yoshihiro Watanabe, and Haruyuki Nakano. “Efficient implementation of relativistic and non-relativistic quasidegenerate perturbation theory with general multiconfigurational reference functions”. In: *Chemical physics letters* 442.1-3 (2007), pp. 164–169.
- [34] Alexander A Granovsky. “Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory”. In: *The Journal of chemical physics* 134.21 (2011).
- [35] Hisao Nakamura and Donald G Truhlar. “The direct calculation of diabatic states based on configurational uniformity”. In: *The Journal of Chemical Physics* 115.22 (2001), pp. 10353–10372.
- [36] Hisao Nakamura and Donald G Truhlar. “Direct diabatization of electronic states by the fourfold way. II. Dynamical correlation and rearrangement processes”. In: *The Journal of chemical physics* 117.12 (2002), pp. 5576–5593.
- [37] Willem Siebrand and Marek Z Zgierski. “Interaction between vibronic coupling and spin—orbit coupling in pyrazine”. In: *Chemical Physics Letters* 67.1 (1979), pp. 13–16.
- [38] Kurt R Glaesemann, Mark S Gordon, and Haruyuki Nakano. “A study of FeCO+ with correlated wavefunctions”. In: *Physical Chemistry Chemical Physics* 1.6 (1999), pp. 967–975.
- [39] Padmabati Mondal et al. “Jahn–Teller and spin–orbit coupling effects in transition-metal trifluorides”. In: *Chemical Physics* 387.1-3 (2011), pp. 56–65.
- [40] Padmabati Mondal et al. “Ab initio study of dynamical E× e Jahn-Teller and spin-orbit coupling effects in the transition-metal trifluorides TiF₃, CrF₃, and NiF₃”. In: *The Journal of Chemical Physics* 136.8 (2012).
- [41] T Leitner et al. “Time-resolved electron spectroscopy for chemical analysis of photodissociation: Photoelectron spectra of Fe (CO) ⁵, Fe (CO) ⁴, and Fe (CO) ³”. In: *The Journal of chemical physics* 149.4 (2018).

- [42] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- [43] Marcel Nooijen. *Vibronic Notes 2*. https://www.sci.uwaterloo.ca/~nooijen/website/vibron/VC/Vibronic_Notes.pdf. 2012.
- [44] Fiorella Villanueva Heldmaier. “Spectroscopic interrogation of differential mobility spectrometry selected, isolated, gas-phase molecular clusters”. M.Sc. Thesis. University of Waterloo, 2020.
- [45] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017.
- [46] Hannah Chang. “From electronic structure theory to molecular spectroscopy”. In: *BA degree Thesis, Princeton University* (2003).
- [47] H Köppel, LS Cederbaum, and W Domcke. “Interplay of Jahn–Teller and pseudo-Jahn–Teller vibronic dynamics in the benzene cation”. In: *The Journal of chemical physics* 89.4 (1988), pp. 2023–2040.
- [48] Michael W Schmidt and Mark S Gordon. “The construction and interpretation of MCSCF wavefunctions”. In: *Annual review of physical chemistry* 49.1 (1998), pp. 233–266.
- [49] Yuki Kobayashi et al. “Direct mapping of curve-crossing dynamics in IBr by attosecond transient absorption spectroscopy”. In: *Science* 365.6448 (2019), pp. 79–83.
- [50] Hans Lischka et al. “Multireference approaches for excited states of molecules”. In: *Chemical reviews* 118.15 (2018), pp. 7293–7361.
- [51] Joseph Ivanic. “Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. I. Method”. In: *The Journal of Chemical Physics* 119.18 (2003), pp. 9364–9376.
- [52] Benjamin Helmich-Paris. “CASSCF linear response calculations for large open-shell molecules”. In: *The Journal of Chemical Physics* 150.17 (2019), p. 174121.

- [53] Ke R Yang, Xuefei Xu, and Donald G Truhlar. “Direct diabaticization of electronic states by the fourfold-way: Including dynamical correlation by multi-configuration quasidenerate perturbation theory with complete active space self-consistent-field diabatic molecular orbitals”. In: *Chemical Physics Letters* 573 (2013), pp. 84–89.
- [54] Cristopher Camacho, Shigeyoshi Yamamoto, and Henryk A Witek. “Choosing a proper complete active space in calculations for transition metal dimers: ground state of Mn 2 revisited”. In: *Physical Chemistry Chemical Physics* 10.33 (2008), pp. 5128–5134.
- [55] Tao Zeng, Dmitri G Fedorov, and Mariusz Klobukowski. “Performance of dynamically weighted multiconfiguration self-consistent field and spin-orbit coupling calculations of diatomic molecules of Group 14 elements”. In: *The Journal of Chemical Physics* 134.2 (2011).
- [56] Dmitri G Fedorov et al. “Spin-orbit coupling in molecules: Chemistry beyond the adiabatic approximation”. In: *International Reviews in Physical Chemistry* 22.3 (2003), pp. 551–592.
- [57] Gary L Miessler. *Inorganic chemistry*. Pearson Education India, 2008.
- [58] Tao Zeng et al. “General formalism of vibronic Hamiltonians for tetrahedral and octahedral systems: problems that involve T, E states and t, e vibrations”. In: *Journal of Chemical Theory and Computation* 13.10 (2017), pp. 5004–5018.
- [59] Isaac Bersuker. *The Jahn-Teller effect and vibronic interactions in modern chemistry*. Springer Science & Business Media, 2013.
- [60] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics*. Cambridge University Press, 2020.
- [61] Masaaki Baba. “Intersystem Crossing in the $1n\pi^*$ and $1\pi\pi^*$ States”. In: *The Journal of Physical Chemistry A* 115.34 (2011), pp. 9514–9519.

- [62] JB Coon, RE DeWames, and CM Loyd. “The Franck-Condon principle and the structures of excited electronic states of molecules”. In: *Journal of Molecular Spectroscopy* 8.1-6 (1962), pp. 285–299.
- [63] Frank Neese. “The ORCA program system”. In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.1 (2012), pp. 73–78.
- [64] H-D Meyer, Uwe Manthe, and Lorenz S Cederbaum. “The multi-configurational time-dependent Hartree approach”. In: *Chemical Physics Letters* 165.1 (1990), pp. 73–78.
- [65] Michael H Beck et al. “The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. Version 8.6, (2023)”. In: *Physics reports* 324.1 (2000), pp. 1–105.
- [66] I Burghardt, K Giri, and GA Worth. “Multimode quantum dynamics using Gaussian wavepackets: The Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine”. In: *The Journal of chemical physics* 129.17 (2008), p. 174104.
- [67] Giuseppe M. J. Barca et al. “Recent developments in the general atomic and molecular electronic structure system”. en. In: *The Journal of Chemical Physics* 152.15 (Apr. 2020), p. 154102. ISSN: 0021-9606, 1089-7690. DOI: [10.1063/5.0005188](https://doi.org/10.1063/5.0005188). URL: <http://aip.scitation.org/doi/10.1063/5.0005188> (visited on 06/18/2020).
- [68] Federico Zahariev et al. “The general atomic and molecular electronic structure system (GAMESS): novel methods on novel architectures”. In: *Journal of Chemical Theory and Computation* 19.20 (2023), pp. 7031–7055.
- [69] Dmitri G Fedorov and James P Finley. “Spin-orbit multireference multistate perturbation theory”. In: *Physical Review A* 64.4 (2001), p. 042502.
- [70] Takahito Nakajima and Kimihiko Hirao. “The higher-order Douglas–Kroll transformation”. In: *The Journal of Chemical Physics* 113.18 (2000), pp. 7786–7789.

- [71] B.J. Chen and N.G. Raymond. *VEModel*. <https://github.com/bjb2chen/VEModel>. 2024.
- [72] Haruyuki Nakano. *Students’ guide for GMCPT/GMCQDPT calculations on GAMESS*. URL: <https://ccl.scc.kyushu-u.ac.jp/~nakano/gmcpt.html>.
- [73] Tao Zeng et al. “Tuning the ground state symmetry of acetylenyl radicals”. In: *ACS central science* 1.5 (2015), pp. 270–278.
- [74] Charles R Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (2020), pp. 357–362.
- [75] Ofer Dekel. *Optimizing Python: Why Python Is Slow and 4 Optimization Methods*. 2010s.
- [76] Python Software Foundation. *mmap — Memory-mapped file support*. 2000.
- [77] Neil Raymond. “Equilibrium and dynamical computation schemes for vibronic models of nonadiabatic systems”. In: *PhD Thesis, University of Waterloo* (2022).
- [78] CR Brundle and David Warren Turner. “High resolution molecular photoelectron spectroscopy II. Water and deuterium oxide”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 307.1488 (1968), pp. 27–36.
- [79] JE Reutt et al. “Molecular beam photoelectron spectroscopy and femtosecond intramolecular dynamics of H₂O⁺ and D₂O⁺”. In: *The Journal of chemical physics* 85.12 (1986), pp. 6928–6939.
- [80] Benny Jun Chen et al. “Adiabatic-to-Diabatic Angle in Seam Space: Renner-Teller-Type and Pseudo-Jahn-Teller-Type Problems”. In: *ChemPhysChem* 25.11 (2024), e202400130.
- [81] GR Branton et al. “The photoelectron spectra of phosphine and arsine”. In: *Chemical Physics Letters* 5.1 (1970), pp. 1–2.

- [82] D Edvardsson et al. “A photoabsorption, photodissociation and photoelectron spectroscopy study of NH₃ and ND₃”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 32.11 (1999), p. 2583.
- [83] Rein Maripuu et al. “The HeI excited electron spectrum of phosphine: An experimental and theoretical study”. In: *Molecular Physics* 48.6 (1983), pp. 1255–1267.
- [84] Swarnendu Bhattacharyya, Zuyang Dai, and Wolfgang Domcke. “Quantum dynamics on a three-sheeted six-dimensional ab initio potential-energy surface of the phosphine cation: Simulation of the photoelectron spectrum and the ultrafast radiationless decay dynamics”. In: *The Journal of Chemical Physics* 143.19 (2015).
- [85] Silver Nyambo, Yuchen Zhang, and Dong-Sheng Yang. “Vibronic transitions and spin–orbit coupling of three-membered metallacycles formed by lanthanide-mediated dehydrogenation of dimethylamine”. In: *The Journal of Chemical Physics* 155.3 (2021), p. 034302.
- [86] Nicholas P Gaba et al. “Simulating transient X-ray photoelectron spectra of Fe (CO)₅ and its photodissociation products with multireference algebraic diagrammatic construction theory”. In: *Physical Chemistry Chemical Physics* 26.22 (2024), pp. 15927–15938.
- [87] Maria Besora et al. “A combined theoretical and experimental study on the role of spin states in the chemistry of Fe (CO)₅ photoproducts”. In: *Journal of the American Chemical Society* 131.10 (2009), pp. 3583–3592.
- [88] A Dominic Fortes and Stewart F Parker. “Structure and spectroscopy of iron pentacarbonyl, Fe (CO)₅”. In: *Journal of the American Chemical Society* 144.38 (2022), pp. 17376–17386.
- [89] Takeshi Noro, Masahiro Sekiya, and Toshikatsu Koga. “Segmented contracted basis sets for atoms H through Xe: Sapporo-(DK)-n ZP sets (n= D, T, Q)”. In: *Theoretical Chemistry Accounts* 131 (2012), pp. 1–8.

- [90] Andrew L Hector et al. “UV-visible spectroscopic studies of group 8–10 metal trifluorides”. In: *Journal of fluorine chemistry* 84.2 (1997), pp. 161–165.
- [91] K Amine et al. “Intercalation of rhodium fluoride into graphite: reaction mechanism in anhydrous HF and physical characterization”. In: *Materials research bulletin* 26.5 (1991), pp. 337–347.
- [92] Christoph Schnedermann et al. “Vibronic Dynamics of the Ultrafast all-trans to 13-cis Photoisomerization of Retinal in Channelrhodopsin-1”. In: *Journal of the American Chemical Society* 138.14 (2016), pp. 4757–4762.
- [93] GR Branton et al. “A Discussion on photoelectron spectroscopy-Photoelectron spectra of some polyatomic molecules”. In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 268.1184 (1970), pp. 77–85.
- [94] Baohua Niu, David A Shirley, and Ying Bai. “High resolution photoelectron spectroscopy and femtosecond intramolecular dynamics of H₂CO⁺ and D₂CO⁺”. In: *The Journal of chemical physics* 98.6 (1993), pp. 4377–4390.
- [95] Brett M Bode and Mark S Gordon. “MacMolPlt: a graphical user interface for GAMESS”. In: *Journal of Molecular Graphics and Modelling* 16.3 (1998), pp. 133–138.

Appendices

Appendix A

A.1 MCSCF Pipeline

Readers can follow along on: <https://youtu.be/KJBxNBxqhHA>.

Inside the repository are other accompanying videos with commentary running through the full fledged protocol from input geometry to spectra. RhF_3 molecule will be used as an example here.

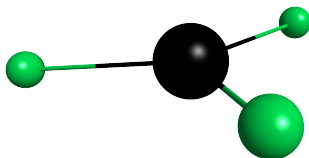


Figure A.1: RhF_3

A.1.1 Step 1: Geometry optimization + frequency (Hessian)

Following standard practice for GAMESS users, input geometry coordinates are first set using the wxMacMolPlt GUI builder tool [95]. Typically the $\$data$ group appears as follows, and GAMESS will duplicate the single fluorine atom given by the appropriate D_{3h} geometry. In the appendices, I have provided a standardized template that can be used as a starting point for all GAMESS calculations in this protocol.

```

1 $data
2 DNH 3
3
4 Rh      45.0      0.0  0.0  0.0
5 F       9.0       2.07 0.0  0.0
6 $end

```

Listing A.1: Example D_{3h} RhF_3 Input Geometry

Important additional flags that are needed to be modified from the default template:

```

1 $contrl runtyp=optimize          # 'geometry optimization'
2 $contrl scfityp=rohff mult=5 icharg=0 # ROHF, multiplicity of 5, neutral
3 $basis gbasis=SPKrtZP          # Sapporo triple zeta basis set
4 $contrl relwfn=LUT-IOTC        # relativistic effects for heavy
    atom
5 $scf diis=.t                   # turn on DIIS SCF converger

```

Listing A.2: Additional flags for Step 1.

Table A.1: Step 1 information.

Input	Output
<code>rungms RhF3_SPK_mp2_rohff_D3h_mult5_diis_gh.inp ncpu ngb nhour</code>	<code>RhF3_*_gh.out</code>

After the input file is successfully setup, one can submit the job to the cluster via the following input command. The resulting output file should locate the important 'EQUILIBRIUM GEOMETRY', and list out the $3N-6 = 6$ vibrational normal modes at the end of calculation.

A.1.2 Step 2: GMCPT

This is the critical MCSCF (ORMAS) active space setup step.

```

1 $contrl runtyp=energy nosym=1 scf=mcscf relwfn=LUT-IOTC
2 $guess norb=168
3 $mcscf cistep=gmcct fullnr=.t. finci=mos
4         acurcy=1.d-6 maxit=200 $end
5 $mrmp mrpt=gmcpt rdvecs=.f. $end
6 $gmcpt nmofzc=21 nmodoc=12 nmoact=5 nelact=6 stsym(1)=A
7         reftyp=ormas nspace=1 mstart(1)=34 mine(1)=6 maxe(1)=6
8         kstate(1)=1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
9         wstate(1)=1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
10        iwgt=0 thrwgt=-1.0 krot=.f. kszdoe=.f. knospn=.t.
11        thrde=-1.0 edshft=0.02 $end
12 $transt $end
13 $data 'EQUILIBRIUM GEOMETRY' coordinates
14 $VEC group of ROHF orbitals

```

Listing A.3: Additional flags for Step 2.

Table A.2: Step 2 information.

Command Line	Input	Output
python3 toolkit.py python3 toolkit.py	RhF3_gh.out RhF3_gh.dat	ref_structure vec.dat (ROHF orbs)
rungms .inp ncpu ngb nhour	RhF3_gmcpt.inp	RhF3_gmcpt.out

GMCPT is tedious to get right. The recommendation is to consult the students' guide for GMCPT/GMC-QDPT calculations prepared by Nakano and frequently run many experimental jobs. One must ensure that they get the combo of frozen core, doubly occupied, and active orbitals/electrons right. Likewise, if the orbital ranking you input is unphysical,

it can potentially result in a poor calculation. If you desire to introduce SOC to your system, attention to add the *\$transt \$end* flag. As it turns out, if you add SOC, and you are running on *nlogn*, the memory mapping fails! This is due to a bug on compiling GMCPT module on *nlogn*, causing memory corrupted bytes to appear near the termination of the calculation. However, do not be alarmed, as all the meaningful SOC information will be printed out before the memory corruption happens in the output file. If you run the exact same calculations on *ComputeCanada* servers, there will be no problem.

A.1.3 Step 3: Diabatic Molecular Orbitals (DMO)

Semi-canonical orbitals to DMO

Here, you will simply use the toolkit script to extract out the Semi-CMOs and *\$DMO* group from the second step's *.dat* file. Be sure you know the active space's *nmoact* and *mstart* values, as the toolkit script will require these inputs to extract out the correct portion of diabatic molecular orbitals from the full semi-canonical group. Step 3's input will be identical to Step 2, however the prior orbitals shall be deleted and in its place concatenate the Semi-canonical orbitals along with *\$DMO* group to the end of the file.

```

1 ~ Step 2 Input ~
2 $VEC group of Semi-canonical orbitals (replacing old ROHF orbitals)
3 $DMO group

```

Listing A.4: Additional flags for Step 3.

A.1.4 Step 4: Reference Determinant

Rerun the previous Step 3's process on the new *\$DMO* outputs to obtain the Semi-canonical MOs and *\$DMO* group again. After, you will run *DiabToolkit* on the *.out* file from Step 3, and from it can extract the *\$REFDET* group automatically. Note that one can actually handcraft

Table A.3: Step 3 information.

Command Line	Input	Output
<code>python3 toolkit.py</code>	<code>RhF3_gmcpt.dat</code>	<code>active_space_orbs</code> (Semi-canonical MOs)
<code>python3 toolkit.py</code>	<code>RhF3_gmcpt.dat</code>	<code>dmo.dat</code> (DMO group)
<code>rungms .inp ncpu ngb nhour</code>	<code>RhF3_dmo.inp</code>	<code>RhF3_dmo.out</code>

a \$REFDET group for better accuracy, using knowledge of dominant Slater determinants and state configurations by reviewing Step 3's output. Once more, concatenate it to the exact same Step 3 input, and this will be the Step 4 input file.

```

1 ~ Step 3 Input ~
2 $VEC group of Semi-canonical orbitals (fresh from Step 3 outputs)
3 $DMO group (fresh from Step 3 outputs)
4 $REFDET group

```

Listing A.5: Additional flags for Step 4.

Table A.4: Step 4 information.

Command Line	Input	Output
<code>python3 toolkit.py</code>	<code>RhF3_dmo.dat</code>	<code>active_space_orbs</code> (Fresh Semi-canonical MOs)
<code>python3 toolkit.py</code>	<code>RhF3_dmo.dat</code>	<code>dmo.dat</code> (DMO group)
<code>rungms .inp ncpu ngb nhour</code>	<code>RhF3_refdet.inp</code>	<code>RhF3_refdet.out</code>

Cleanup: `temp.inp`

After Step 4 terminates successfully, there should be full vibronic coupling information results. If this is the case, then the diabaticization template `temp.inp` shall be prepared from

Step 4's input. The only difference is that the `$data` group must be deleted, and a new `$data` section appended to the end of the template.

```
1 ~ Step 4 Input ~  
2 ~ early $data group deleted ~  
3 $data  
4 title of diabatisation  
5 C1
```

Listing A.6: Additional flags for Step 5.

A.2 Formaldehyde, CH₂O

The formaldehyde model has 5 electronic surfaces and 6 normal modes. Active space used for this CH₂O system was 6o7e. Special custom TDM moments along the principal Z-axis were attempted to be calculated for one version of this model. Theoretical simulated photoelectron spectra of CH₂O shown in Figure A.2 and Figure A.3 display excellent agreement with the experimental CH₂O spectrum recorded by Niu et al. [94]. The three distinct absorption bands located at 10.5 eV, 14-15.5 eV, and 15.5-18 eV are reproduced in all both MCTDH and VECC approaches to a high degree of resolution. Within the high energy vicinity, the VECC method shows slight blue-shifting in contrast to the MCTDH method.

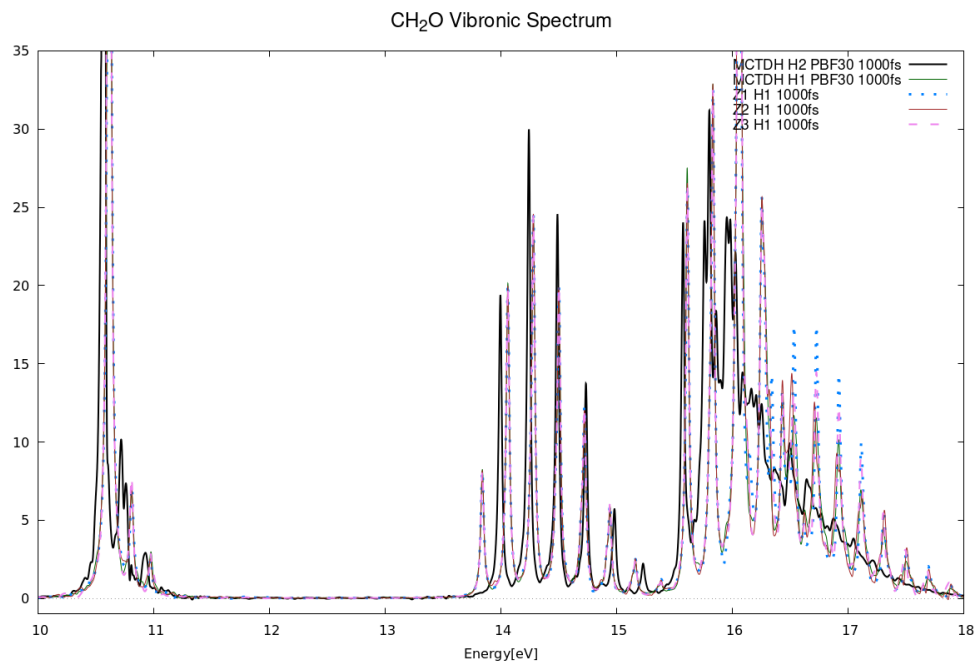


Figure A.2: Theoretical vibrationally-resolved photoelectron spectrum of CH_2O .

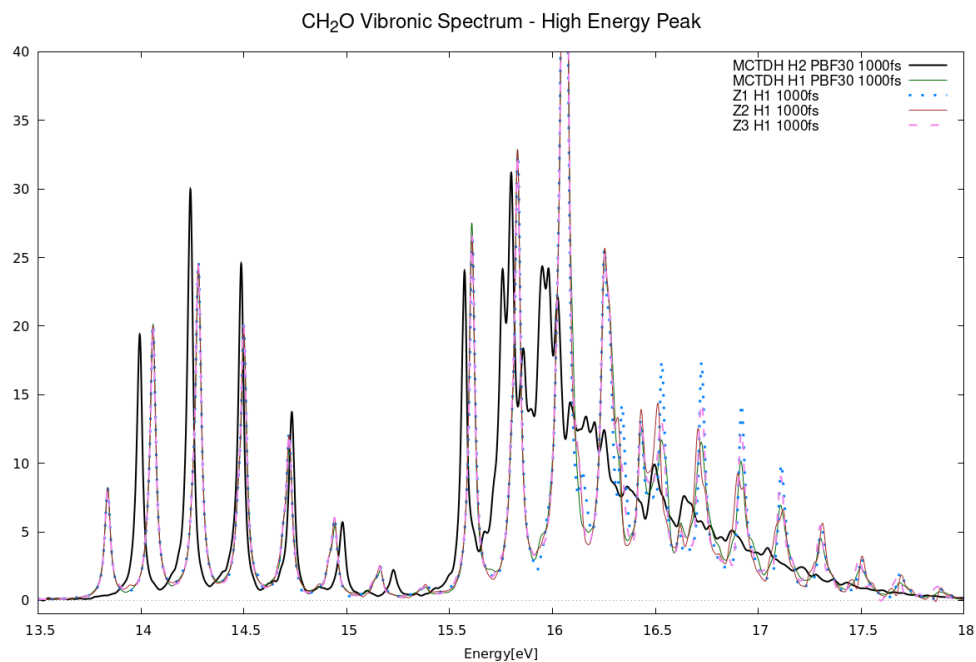


Figure A.3: Theoretical vibrationally-resolved photoelectron spectrum of CH_2O .

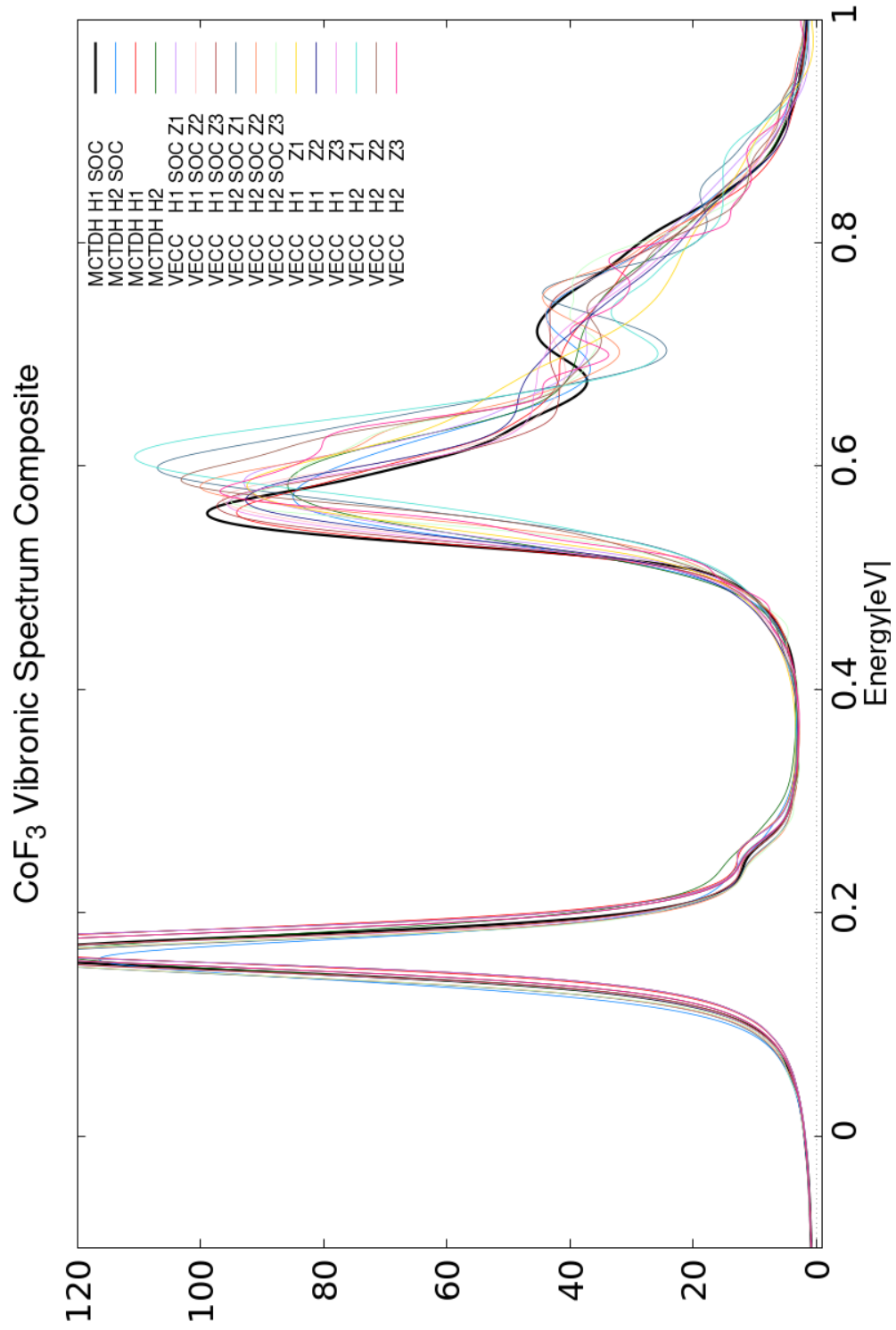


Figure A.4: Composite of theoretical vibrationally-resolved spectrum of CoF₃.

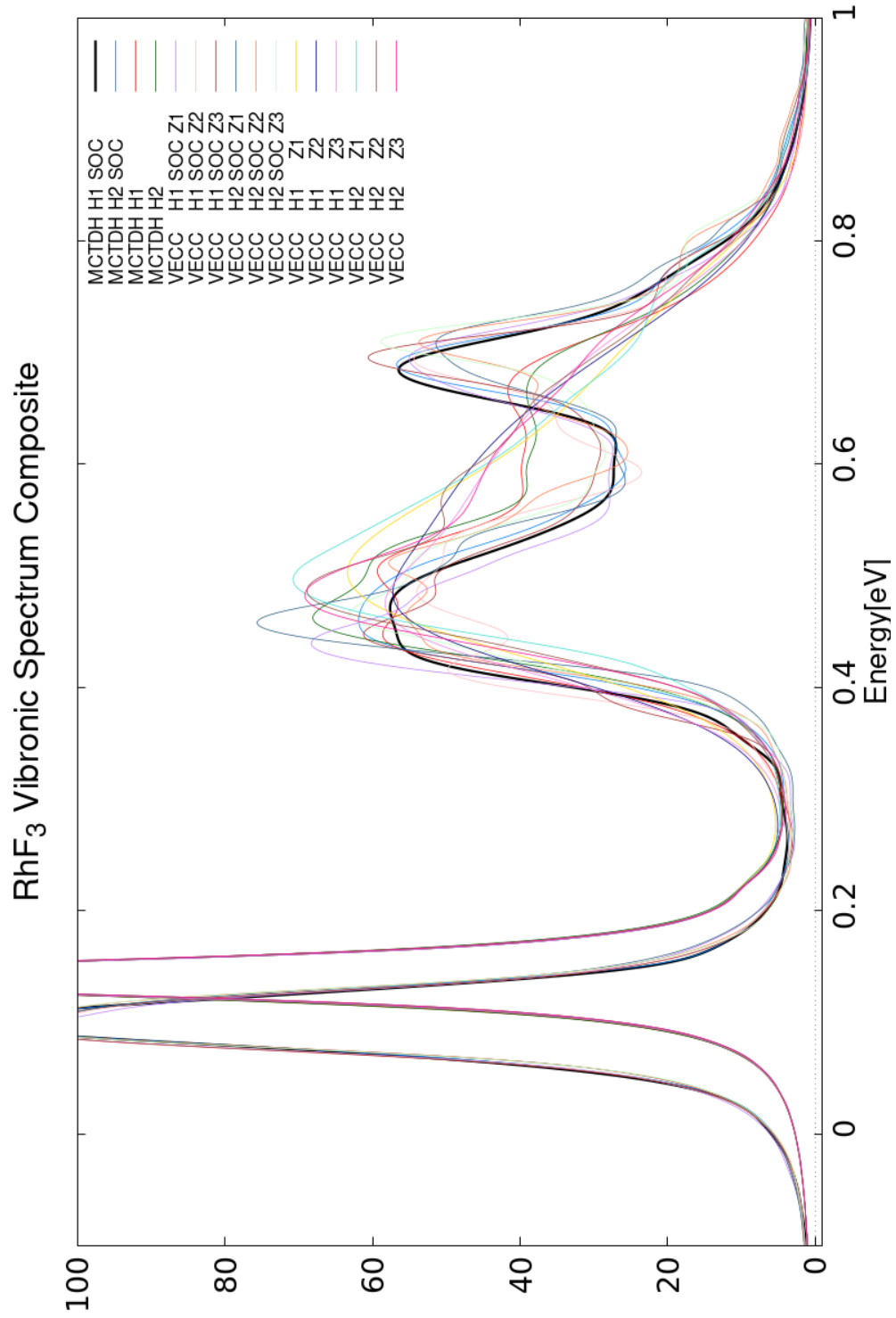


Figure A.5: Composite of theoretical vibrationally-resolved spectrum of RhF₃.

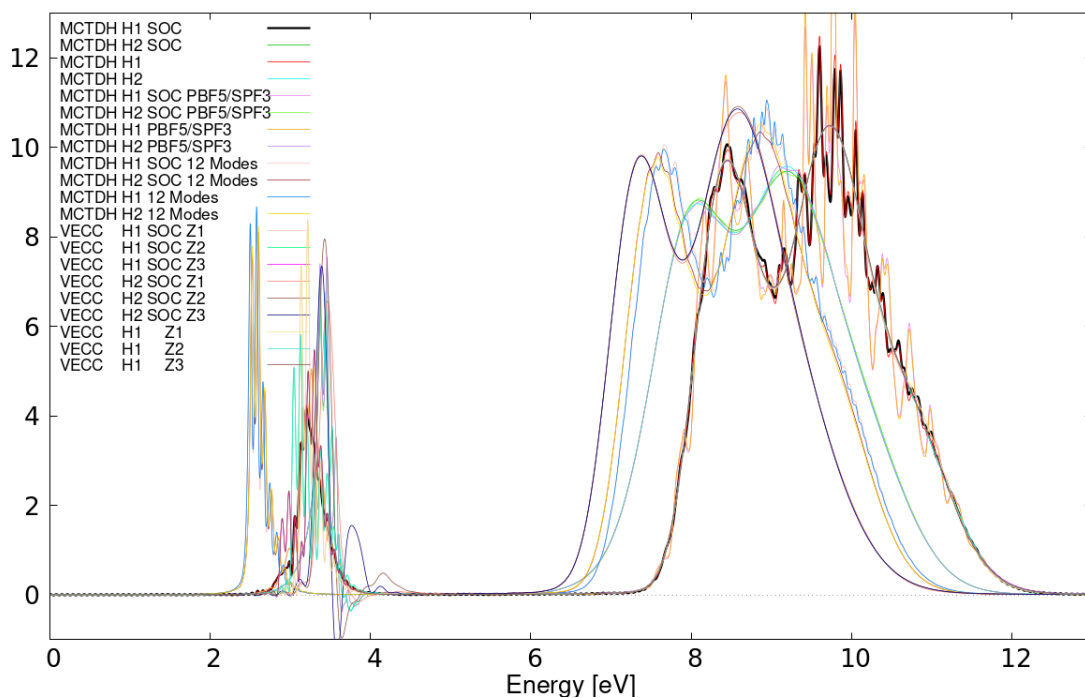


Figure A.6: Theoretical vibrationally-resolved spectrum composite of $\text{Fe}(\text{CO})_5$. Propagated for 180fs and resolved with $\tau = 40$.

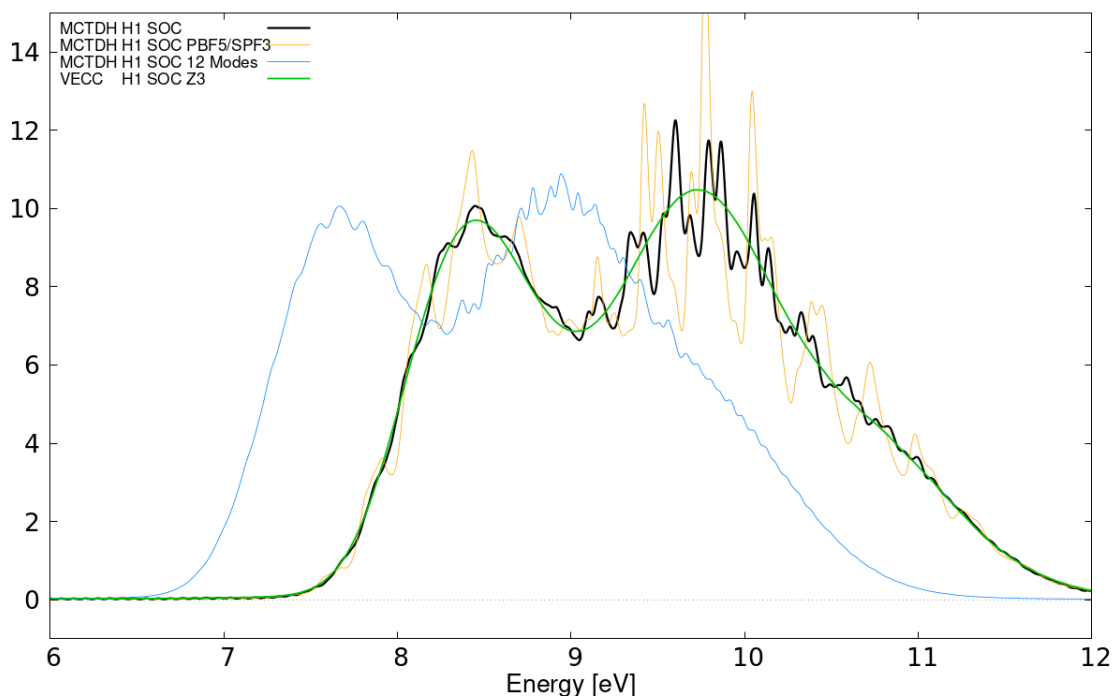


Figure A.7: Theoretical vibrationally-resolved spectrum of $\text{Fe}(\text{CO})_5$ between 6–12 eV. Propagated for 180fs and resolved with $\tau = 40$. This region is quite sensitive to the contribution from varying parameters such as basis function size and restricted modes.