UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Applying Time-dependent Density Functional Theory's Linear Response to π molecular systems for electronic excitation properties

Permalink

https://escholarship.org/uc/item/6gq8x4hn

Author

Mohanam, Luke Nambi

Publication Date

2021

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org

Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Applying Time-dependent Density Functional Theory's Linear Response to π molecular systems for electronic excitation properties

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Luke Nambi Mohanam

Dissertation Committee:
Professor Filipp Furche, Chair
Distinguished Professor Kieron Burke
Professor Craig Martens

2021

Portions of Chapter 4 © 2015 & 2018 American Chemical Society
Chapter 5 © 2020 American Chemical Society
All other materials © 2021 Luke Nambi Mohanam

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The work in this thesis has been supported by the National Science Foundation under OAC-1835909 and under CHE-1800431, and the Department of Energy under under DE-SC0018352.

This dissertation contains verbatim excerpts as they appear in J. Am. Chem. Soc., and J. Chem. Educ.. Permission to use copyrighted material in the dissertation has been grant by 2018 & 2020 American Chemical Society. The co-authors listed in these publications directed, supervised, or conducted part of the research which forms the basis for the dissertation. The list of co-authors, as they appear, are Tener F. Jenkins, David H. Woen, Joseph W. Ziller, Filipp Furche, William J. Evans and Amanda J. Holton.

Central to my graduate career is Filipp Furche, who has been an excellent teacher, mentor in many things, research group leader, software lead, supervisor, and friend. I am most thankful for his belief and trust in me, which I will strive to live up to. Dmitrij Rappoport has stepped in to assisting with my advising, and his perspective and patience have been invaluable.

Amanda Holton has been an excellent teacher, teaching mentor, teaching supervisor, and friend. She has carried out the role of a Principal Investigator for my teaching project, and I am deeply grateful for this kindness and support.

For the material in this thesis, I have closely collaborated scientifically with Filipp Furche, Dmitrij Rappoport, Samuel Bekoe, Zoey Shen Ziyue, Naje' George, Noah Bibo Feng, Tener Jenkins, William Evans, and Amanda Holton. In addition, the research is supported by the TURBOMOLE development team, where I have mainly interacted with Uwe Huniar and Yannick Franzke; and software infrastructure specialists at UCI, where I have mainly interacted with Nadya Williams and Nathan Crawford. The debugging of Tully Surface hopping relied on discussions with and data from Saswata Roy, Shane Parker, Gabriel Phun, and Anthony Lara.

Beyond the people directly relevant to this work, the past six years would not have been possible without the support and kindness of many people towards me, often fulfilling multiple roles in different capacities. Rather than typing out a long list of people, in which some important names will be left out on accident, I focus on the groups and organizations within UCI and her affilates.

A university is led by her faculty. The faculty of the department of chemistry has been friendly, helpful, instructive and a pleasure to be around. Besides being on my defense and advancement committee, the graduate classes have provided a bedrock to ground my research.

The senior research staff are key for sharing research experience. I have been fortunate to receive guidance and support from a range of post-doctoral associates. Their advice and perspective are greatly appreciated. In particular I have been exposed to more software development aspects which may be important for my future career.

The research work and leading discussion sessions in UCI is primarily done by graduate students. I have been surrounded by supportive, intelligent, and driven graduate students. They have done much to show me Irvine, the culture and the surroundings. Especially driving me to places I would not reach on foot.

Undergraduates are an important part of a university. I have been fortunate to collaborate with and instruct bright and hopeful undergraduates in research and chemistry. In particular, I have been on the board of the Swing Dance @ UCI club which has mainly undergraduate members. I appreciate the club for teaching me to dance and for perspectives outside of graduate school.

High school interns are part of University Outreach. I am thankful for the opportunity to interact with motivated and curious high school students.

Encompassing all these groups is the Furche Research Group. Keeping research rolling forward requires maintainance of tools and resources, and the group does so. In particular, chemistry is a broad subject, and listening and hearing about other projects has been important to my development.

Outside of UCI, I am happy to have spent the last few years together with my girlfriend Kayla. It is brilliant to have someone to discuss chemistry and share my life with.

VITA

Luke Nambi Mohanam

EDUCATION

Doctor of Philosophy in Chemistry **2021**
University of California *Irvine, California*

Bachelor of Science (First Class Honours), in Chemistry **2014**
National University of Singapore *Singapore, Singapore*

RESEARCH EXPERIENCE

Graduate Research Assistant **2015–2021**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2015–2019**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Gamifying intermolecular forces: Improving student performance outcomes for college freshman-level intermolecular forces **2020**
Journal of Chemical Education

Tetramethylcyclopentadienyl Ligands Allow Isolation of Ln(II) Ions across the Lanthanide Series in [K(2.2.2-cryptand)] [(C_5Me_4H)$_3$Ln] Complexes **2018**
Organometallics

That Little Extra Kick: Nonadiabatic Effects in Acetaldehyde Photodissociation **2016**
The Journal of Physical Chemistry Letters

Effect of orbital and ionic dynamics coupling in barrier crossing rates for Car–Parrinello molecular dynamics **2015**
Chemical Physics Letters

SOFTWARE

libkrylov `https://gitlab.com/libkrylov`
Developer, software library project for Krylov subspace methods

Turbomole `https://www.turbomole.org`
Developer, quantum chemistry package

ABSTRACT OF THE DISSERTATION

Applying Time-dependent Density Functional Theory's Linear Response to π molecular systems for electronic excitation properties

By

Luke Nambi Mohanam

Doctor of Philosophy in Chemistry

University of California, Irvine, 2021

Professor Filipp Furche, Chair

Predicting excited state molecular properties of an ensemble of molecules *in silico* is a goal for theoretical and computational chemists. Solving for these properties with exact and analytical physics is out of reach within a reasonable time-frame, hence approximations are taken. This thesis starts from the approximations of Kohn Sham DFT to determine a reference electronic state for a molecular geometry of classical nuclei, then computing excited electronic states and properties such as gradients using response theory in the framework of TDDFT. While there are well established sets of parameters and approximations that have worked well on particular molecular systems, this thesis presents an investigation of parameters relating to the (1) Krylov subspace approximate solve for excitations, (2) Tully Surface Hopping algorithm for non-adiabatic molecular dynamics, and (3) Excitations of Lanthanide(II) complexes where previous, standard approaches are unsatisfactory. The broader impact is a better understanding of these approaches and parameters, with attention drawn to various algorithmic features that should color interpretation of computed properties for most molecular systems.

In pursuit of understanding chemistry, the chemical sense of intermolecular forces (important to property predictions described above), has been translated into (4) a card game at the undergraduate general chemistry level. A pedagogy study of the card game was carried out, suggesting that it could improve undergraduate understanding of intermolecular forces.

Chapter 1

Krylov subspace algorithm for approximately solving Linear Response equations

The material in this chapter is based upon work supported by the National Science Foundation under OAC-1835909.

libkrylov will be available at

https://gitlab.com/libkrylov/

1.1 Abstract

Having obtained a reference, ground state electronic density matrix $\boldsymbol{\gamma}_0$ from non-relativistic Kohn-Sham Density Functional Theory(KS-DFT), linear response and other perturbations can be carried out in the framework of Time-Dependent DFT (TDDFT) to give linear

response equations.[1, 2] These linear response equations are large eigenvalue and linear equations which are typically not solved directly. Solving these linear response equations usually require a Krylov subspace approach which is typically hard coded around the different evaluations of the linear response equations.[3]

After introducing the linear response equations and a generic Krylov subspace algorithm, a current implementation in TURBOMOLE[4] is described briefly to point out the status quo of hard coding the Krylov subspace algorithm.

To generalize and modularize the algorithm and its implementation in an open-source software library is a goal of the *libkrylov* project, led by Rappoport and co-workers.[5] Doing so allows for standardization of the Krylov Subspace solve within *libkrylov* to be shared and be comparable between different implementations and problems that use *libkrylov*. *Libkrylov* for Krylov subspace solves can be compared to reference BLAS and LApack libraries[6] for diagonalizing matrices, such as `dsyev`.

Krylov subspace algorithms are known to lack a "one-size-fits-all" definition, [7] with options for the subspace projection [3] and preconditioning.[8, 9] For solving more varied and difficult linear response problems, making these options user-selectable and characterizing how these options interact (in the framework of *libkrylov*) is the step taken in this chapter.

1.2 Introduction

1.2.1 Linear Response equations

A quick summary of the derivation of the linear response equations, present in the literature,[1] follows.

A time-independent reference ground state density matrix, $\boldsymbol{\gamma}_0$ is constructed in the basis of canonical time-independent Kohn-Sham orbitals ϕ_p,

$$\boldsymbol{\gamma}_0 = \sum_p n_p \left|\phi_p\right\rangle \left\langle\phi_p\right| \tag{1.1}$$

where n_p is the occupation number, turning p into summation over all occupied orbitals.

The canonical Kohn-Sham orbitals satisfy their time-independent Schrödinger equation:

$$\mathbf{H}^{(0)}[\boldsymbol{\gamma}_0] \left|\phi_p\right\rangle = \epsilon_p \left|\phi_p\right\rangle \tag{1.2}$$

where the use of $\boldsymbol{\gamma}_0$ allows the equation to be solved self-consistently. ϵ_p are the Kohn Sham orbital energies, $\mathbf{H}^{(0)}[\boldsymbol{\gamma}]$ is the time independent Hamiltonian which is a functional of a density matrix $\boldsymbol{\gamma}$.

In practice, the time-independent Hamiltonian contains an approximate exchange-correlation potential which is obtained as a functional derivative of the approximate exchange-correlation energy functional with respect to the density. Appropriate energy functionals are chosen for chemical systems to be studied. Various approximations are well discussed in the literature.

For a single small monochromatic perturbation Ω_α (index is α), a time-dependent Hamiltonian can be written with perturbation theory, with a Taylor series about small factor λ_α:

$$
\begin{aligned}
\mathbf{H}[\boldsymbol{\gamma}](t) &= \mathbf{H}^{(0)}[\boldsymbol{\gamma}(t)] + \lambda_\alpha \mathbf{H}^{(\alpha)}(\Omega_\alpha, t) \\
&= \mathbf{H}^{(0)}[\boldsymbol{\gamma}(t)] + \lambda_\alpha \left(v^{(\alpha)}(\Omega_\alpha) e^{i\Omega_\alpha t} + v^{(\alpha)}(-\Omega_\alpha) e^{-i\Omega_\alpha t} \right) \qquad (1.3) \\
\boldsymbol{\gamma}(t) &= \boldsymbol{\gamma}_0 + \lambda_\alpha \boldsymbol{\gamma}^{(\alpha)}(t) + \frac{1}{2}\lambda_\alpha^2 \boldsymbol{\gamma}^{(\alpha^2)}(t) + \mathcal{O}(\lambda_\alpha^3) \qquad (1.4)
\end{aligned}
$$

The time dependent density matrix satisfies the following time dependent von Neumann equation and an idempotency condition:

$$
\begin{aligned}
i\frac{\partial \boldsymbol{\gamma}(t)}{\partial t} &= \mathbf{H}[\boldsymbol{\gamma}](t)\boldsymbol{\gamma}(t) - \boldsymbol{\gamma}(t)\mathbf{H}[\boldsymbol{\gamma}](t) \qquad (1.5) \\
\boldsymbol{\gamma}(t) &= \boldsymbol{\gamma}(t)\boldsymbol{\gamma}(t) \qquad (1.6)
\end{aligned}
$$

By substituting the Taylor expansion into the above equations, taking a Fourier transform, gathering the terms which are first order in λ_α, the linear response equations are obtained.

$$
\begin{aligned}
\Omega_\alpha \, \boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha) &= \mathbf{H}^{(0)}[\boldsymbol{\gamma}_0] \, \boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha) - \boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha) \, \mathbf{H}^{(0)}[\boldsymbol{\gamma}_0] \\
&\quad + \mathbf{H}^{(0)}[\boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha)] \, \boldsymbol{\gamma}_0 - \boldsymbol{\gamma}_0 \, \mathbf{H}^{(0)}[\boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha)] \\
&\quad + \mathbf{H}^{(\alpha)}(\Omega_\alpha) \, \boldsymbol{\gamma}_0 - \boldsymbol{\gamma}_0 \, \mathbf{H}^{(\alpha)}(\Omega_\alpha) \qquad (1.7)
\end{aligned}
$$

and choosing a parameterization of the linear response in the basis of Canonical KS molecular orbitals ϕ_p (occupied indexed by i, j, virtual indexed by a, b):

$$\boldsymbol{\gamma}^{(\alpha)}(\Omega_\alpha) = \sum_{ia} X_{ia}(\Omega_\alpha)\left|\phi_a\right\rangle\left\langle\phi_i\right| + Y_{ia}(\Omega_\alpha)\left|\phi_i\right\rangle\left\langle\phi_a\right| = \begin{pmatrix} \mathbf{0} & \mathbf{Y}^T(\Omega_\alpha) \\ \mathbf{X}(\Omega_\alpha) & \mathbf{0} \end{pmatrix} \tag{1.8}$$

the resulting linear response equation for several Ω_α (defining matrix $\mathbf{\Omega}$ as a diagonal matrix with diagonal elements Ω_α, and concatanating vectors $\mathbf{X}(\Omega_\alpha)$ and $\mathbf{Y}(\Omega_\alpha)$ into matrices $\mathbf{X}$ and $\mathbf{Y}$:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix}\begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{X} \end{pmatrix} - \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{X} \end{pmatrix}\begin{pmatrix} \mathbf{\Omega} & \mathbf{0} \\ \mathbf{0} & -\mathbf{\Omega} \end{pmatrix} = -\begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q} & \mathbf{P} \end{pmatrix} \tag{1.9}$$

where the terms have the definition (leaving out magnetic contributions), obtained by chain rule and further functional derivatives:

$$(A+B)_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + 2f^{\text{xc}}_{ia,jb} + 2\,(ia|jb) - c_x\left[(ib|ja) + (ij|ab)\right] \tag{1.10}$$

$$(A-B)_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + c_x\left[(ib|ja) - (ij|ab)\right] \tag{1.11}$$

$$P_{ia}(\Omega_\alpha) = Q_{ia}(\Omega_\alpha) = v_{ia}(\Omega_\alpha) \tag{1.12}$$

δ is the Dirac delta function, c_x is the factor for mixing in exact exchange (which is zero for a pure functional), the four center integral $(pq|rs)$ is defined below in chemist notation and real space, and $f^{\text{xc}}_{ia,jb}$ is the kernel, a second order derivative of the approximate energy functional with respect to the density. No new results are presented for deriving the kernel in this thesis, with the standard approach of taking the adiabatic approximation being used.

$$(pq|rs) = \int\int d\mathbf{r}d\mathbf{r}'\delta_{\sigma_p\sigma_q}\delta_{\sigma_r\sigma_s}\phi_p^*(\mathbf{r},\sigma_p)\phi_q(\mathbf{r},\sigma_q)\frac{1}{|\mathbf{r}-\mathbf{r}'|}\phi_r^*(\mathbf{r}',\sigma_r)\phi_q(\mathbf{r}',\sigma_s) \tag{1.13}$$

$\mathbf{A}$ and $\mathbf{B}$ are orbital rotation hessians, and independent of Ω_α. In considering all possible pairs of occupied-virtual orbital rotations, the size of the full basis of $\mathbf{A}$, n, is the number of occupied orbitals multiplied by the number of virtual orbitals. For a closed shell reference state and spin symmetry conserving singlet excitations, this corresponds to the number of spatial orbital basis functions minus half the number of electrons, multiplied by half the number of electrons. For an unrestricted spin reference state, it would be the twice the number of spatial orbital basis functions minus the number of electrons, multiplied by the number of electrons.

$\mathbf{P}$ and $\mathbf{Q}$ represent the perturbation, and are typically referred to as the Right Hand Side (RHS) of a damped response calculation.

$\boldsymbol{\Omega}$ represents the frequencies of the perturbation. Writing the problem as an eigenvalue problem (setting the RHS to zero, in the zero field limit) will correspond the Ω_α eigenvalues to excitation energies. For the symmetric definition of the coefficient matrices, $\boldsymbol{\Omega}$ is real.

The form of this equation is that of a kind of Sylvester equation known as a shifted linear equation. If the reference state is the lowest possible energy state, all excitation energies are positive – negative excitations or instabilities of the full problem are well described in the literature, indicating problems with the reference state. [10] With complicated many electron systems in real 3D space, the possibility of accidental degeneracy or near-degeneracy of two eigenvalues becomes high, which influences the search for eigenvectors.[11]

Typically, for these physical applications, there is an interest in only a small number of solutions, $p << n$, in particular a small number of eigenvalue solutions within a given energy range. For electronic response, there are interior eigenvalue searches for core-valence excitations in the X-ray energy region,[12] but more typically there are lowest eigenvalue searches, in the UV-visible energy region. In both cases, given an energy range of interest, $p > 1$. For certain applications, such as using an adiabatic basis of electronic states calculated via excitations for non-adiabatic molecular dynamics,[13] it is important to obtain not just approximate eigenvectors and eigenvalues with well defined error bounds, but also to get the right order of eigenvectors by eigenvalue magnitudes.

Another commonality to these applications is a desired accuracy of the solutions. If using natural units for energy, the required accuracy of the eigenvalues is about 10^{-5} for the natural units of this calculation (Hartrees), corresponding to an order of magnitude in reaction rates. The lower accuracy required allows for approximate solves of the large problem.

However, for the stability and reliability of the approximate solve (in particular, to obtain the correct ordering of eigenvectors), it is important to clearly define the approximate solve algorithm and validate that the approximate solve is able to approach the accuracy of the direct solve. It is expected that an implemented approximate solve does not reach the accuracy of a direct solve, due to additional numerical operations at finite precision of the approximate solve. In the case that maximum accuracy is desired, a direct solve may be attempted.

For this chapter we present results only for the eigenvalue problem, but consider the general algorithm when discussing the feature of the Krylov Subspace algorithm.

It is possible to consider $\mathbf{X}$ and $\mathbf{Y}$ together as a symplectic super vector - the entire problem has symplectic symmetry. Besides that, there is spin-symmetry if only considering the spin-conserving singlet excitation from a closed shell reference - considering alpha and beta spin electrons together.

For this chapter, a further simplification is taken, known as the Tamm-Dancoff Approximation(TDA) or Configuration Interaction (Single excitations) Approximation [14], setting $\mathbf{Y} = \mathbf{0}$, resulting in the simplified equation:

$$\mathbf{A}\mathbf{X} - \mathbf{X}\boldsymbol{\Omega} = -\mathbf{P} \tag{1.14}$$

The result is a symmetric coefficient matrix $\mathbf{A}$. Other than preserving the form of the full space problem in projection into the subspace, tackling the symmetry of the full space problem is beyond the scope of this chapter.

Molecular spatial symmetry could also be used to simplify the problem further, as the symmetry leads to a block structure of $\mathbf{A}$. Block structure can help simplify and improve the stability of the calculation. However, some perturbations and responses break symmetry, requiring treatment of the full problem.

As can be seen from the definition of $\mathbf{A}$ and $\mathbf{B}$, the coefficient matrices are dense, capturing all possible energy interactions between electron-hole–electron-hole pairs, and are expected to be diagonally dominant, since orbital energy differences $(\epsilon_a - \epsilon_i)$ are expected to be larger than the rest of the terms.

This diagonal dominance allows for the definition of a zeroth order problem and solution(leaving out the arguments $\Omega_\alpha^{[0]}$ from $\boldsymbol{\gamma}^{(\alpha)[0]}$):

$$\Omega_\alpha^{[0]}\,\boldsymbol{\gamma}^{(\alpha)[0]} = \mathbf{H}^{(0)}[\boldsymbol{\gamma}_0]\,\boldsymbol{\gamma}^{(\alpha)[0]} - \boldsymbol{\gamma}^{(\alpha)[0]}\,\mathbf{H}^{(0)}[\boldsymbol{\gamma}_0] \tag{1.15}$$

$$A_{ia,jb}^{[0]} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab}, B_{ia,jb}^{[0]} = 0, P_{ia}^{[0]} = Q_{ia}^{[0]} = 0 \tag{1.16}$$

where the coefficient matrix $\mathbf{A}^{[0]}$ is diagonal, thus the eigenvalues of the zeroth order problem are $(\epsilon_a - \epsilon_i)$ and the eigenvectors are pure KS orbital hole pairs; unit vectors of the basis chosen.

There are thus two options for diagonal matrices which approximate $\mathbf{A}$: (1) the diagonals of matrix $\mathbf{A}$ which can be denoted as matrix $\mathbf{D}$ and (2) the matrix $\mathbf{A}^{[0]}$. It is natural to expect that the better approximation will result in better convergence behavior for an algorithm using that approximation.

Besides the density of matrix elements, the size of the problem is also large, requiring significant computational power to compute and memory to store the entire problem. A general or naive implementation for directly solving a problem with a $(n \times n)$ coefficient matrix would require $4n$ memory and scale as $\mathcal{O}(n^3)$. For large enough n, a direct solve may be impossible. While the usefulness of a Krylov Subspace approach is greatest for systems that cannot be stored or computed with easily available resources, it is important to validate the algorithm on smaller problems whose direct solutions can be directly compared with the approximate solutions. A matrix $\mathbf{A}$, corresponding to the electronic response of Uracil using small atom-centered Gaussian basis sets, is printed to a JSON dictionary to test the present implementation of *libkrylov*.

While this small example is not capable of demonstrating the frequency of accidental degeneracies for more complex systems, in printing out the coefficient matrix, it is possible to separate details of the evaluation of the orbital rotation hessian from the approximate solve of the full problem.

1.2.2 Eigenvalue problem Notation and Lagrangian

How the notation can be generalized is found in appendix A.1.

Focusing on the eigenvalue problem, which can be written as:

$$\mathbf{A}\mathbf{X} - \mathbf{X}\boldsymbol{\Omega} = \acute{\mathbf{A}}\mathbf{X} = \hat{\mathbf{A}}\mathbf{X} = \mathbf{0}$$

where the desired solution is projected by the $\hat{\mathbf{A}}$ operator onto the zero vector in the same space.

Writing the Lagrangian of the full problem[3]:

$$\mathcal{L}[\mathbf{X}, \boldsymbol{\Omega}] = \mathbf{X}^{\dagger}\mathbf{A}\mathbf{X} - \mathbf{X}^{\dagger}\mathbf{X}\boldsymbol{\Omega}$$

The Lagrangian is minimized by the desired solution vector, in other words, for the desired solution $\mathbf{X}$:

$$\left.\frac{\delta\mathcal{L}[\mathbf{V}]}{\delta\mathbf{V}}\right|_{\mathbf{V}=\mathbf{X}} = \mathbf{R}(\mathbf{X}) = \hat{\mathbf{A}}\mathbf{X} = \mathbf{0} \tag{1.17}$$

The gradient of the Lagrangian, the residual, is non-zero for a trial solution or vector that overlaps with the desired analytical solution.

By the Petrov-Galekin condition,[15] $\mathbf{R}^{\dagger}\mathbf{X} = \mathbf{0}$, making it useful to attempt to improve the solutions iteratively with $\mathbf{R}$, leading to Krylov subspace methods.

1.2.3 Krylov Subspace method overview

A Krylov subspace is defined as[15]:

$$\mathcal{K}_k(\mathbf{A}, \mathbf{V}) = \mathrm{span}(\mathbf{V}, \mathbf{A}\mathbf{V}, \mathbf{A}^2\mathbf{V}, ..., \mathbf{A}^{k-1}\mathbf{V}) \tag{1.18}$$

where the subspace is guaranteed to be congruent with the full space upon reaching the size of the full space. (But only when it reaches the size of the full space.)

Matrix Vector Products

The matrix vector product **AV** present an alternative to storing the entire coefficient matrix in memory or on file, and is used in TURBOMOLE's hard-coded implementation. The coefficient matrix elements are computed on-the-fly and multiplied with an arbitrary vector by a matrix-vector product function. This allows for the algorithm to run on computer systems where the system memory size or file read speeds are significant bottlenecks. [3] For the linear response equations, calculating four-center integrals and differentials on a grid can be sped up by screening techniques and insertion of an auxillary basis (typically referred to as a resolution of identity approximation).[16] The consequences of such approximations were previously convoluted with the Krylov subspace algorithm, in particular since the zeroth order term was not screened and directly multiplied. This chapter with *libkrylov* eliminates all such screening. The comparison of the orthogonal and non-orthogonal subspace approach convoluted by the resolution of identity approximation has been reported in the literature.[3]

For the problems where multiple products can be required simultaneously, the matrix vector product function is more efficient if all products are computed as once (even if the vectors and products are on orthogonal spaces among themselves, in a blocked structure). This is particularly in the case where the coefficient matrix is much simpler to calculate in a different basis than the basis used for the full space. For computing the orbital rotation hessians, the Cartesian Gaussian atomic orbital basis is much more straightforward. The matrix-vector product typically becomes the bottleneck of the calculation.

Generic Krylov Subspace Algorithm

Iterative Krylov subspace algorithms are defined by a Krylov Subspace at every iteration k[15, 3]:

$$\mathcal{K}_k(\mathbf{A}, \mathbf{V}^{[1]}) = \mathrm{span}(\mathbf{V}^{[1]}, \mathbf{A}\mathbf{V}^{[1]}, \mathbf{A}^2\mathbf{V}^{[1]}, ..., \mathbf{A}^{k-1}\mathbf{V}^{[1]}) \tag{1.19}$$

The dimensions of $\mathbf{V}^{[k]}$ are $n \times q^{[k]}$. These subspace projection vectors still have one dimension on the full space.

As suggested by the way the definition is written, $\mathbf{V}^{[1]}$ is important in the definition of the Krylov subspace, and are referred to as the start vectors. The intention is to have $\mathbf{V}^{[1]}$ with significant overlap with the desired solutions. For *libkrylov*, $\mathbf{V}^{[1]}$ are explicitly passed in, thus the differences reported in determining $\mathbf{V}^{[1]}$ are differences from passing in different $\mathbf{V}^{[1]}$ to *libkrylov*.

For a diagonally dominant matrix, the diagonal values approximate the eigenvalues of problem, so (for example) the smallest diagonal values correspond to unit vectors that overlap significantly with the lowest eigenvectors, making those unit vectors ideal for $\mathbf{V}^{[1]}$. The diagonals of $\mathbf{A}$, $\mathbf{D}$, are typically not computed. As $\mathbf{A}$ is explicitly calculated for this test application, $\mathbf{D}$ is available and can be used directly to determine starting vectors. In typical applications, $\mathbf{A}^{[0]}$ is used in place of $\mathbf{D}$, and that is also tested.

The algorithm can be described as projecting the large problem onto a smaller subspace for solving, using projection vectors $\mathbf{V}^{[k]}$ which are incremented with every iteration k

A general overview is first presented, with details discussed later in the chapter:

1. If projection is non-orthonormal, prepare orthonormalization scheme.
2. Compute matrix vector products $\mathbf{A}\mathbf{V}^{[k]}$, reuse previous products if possible.
3. Use products to evaluate projection of problem onto subspace: $\hat{\mathbf{a}}^{[k]} = (\mathbf{V}^{[k]})^{\dagger}\hat{\mathbf{A}}\mathbf{V}^{[k]}$
4. If projection is non-orthonormal, apply orthonormalization *on the subspace* to obtain orthonormal problem $\bar{\hat{\mathbf{a}}}\bar{\mathbf{x}}^{[k]} = \mathbf{0}$
5. Solve orthonormalized subspace problem directly and project back onto the full space to obtain best approximate solutions $\mathbf{X}^{[k]}$ (and best approximate eigenvalues $\boldsymbol{\Omega}^{[k]}$ if this is an eigenvalue problem)
6. Obtain residuals as measure of distance from the exact solution $\mathbf{R} = \hat{\mathbf{A}}\mathbf{X}^{[k]}$ (usually using the products, approximate eigenvalues $\boldsymbol{\Omega}^{[k]}$ and some subspace solution $\mathbf{x}^{[k]}$)
7. If Frobenius norm of residuals (the residual norm) is below a convergence threshold, exit. Otherwise,
8. Determine best new vectors $\mathbf{T}$ to expand the projection, typically by preconditioning residuals: $\mathbf{V}^{[k+1]} = \text{concatanate}(\mathbf{V}^{[k]}, \mathbf{T}^{[k]})$

It is known that the algorithm performs well on diagonally dominant coefficient matrices, in particular for selecting good starting vectors and using the diagonal of the matrix for preconditioning.

For simpler notation, the iteration index $[k]$ will be dropped - it is usually clear that all variables belong to the same iteration, and the eigenvalues and solutions of the full space can be assumed to be inaccessible if attempting a Krylov subspace solve. The iteration index will be used if needed when comparing to the solutions of the direct solve which do not have an index.

Cost of each iteration

Given that the matrix vector products remain the bottleneck of this algorithm, stating the cost of each iteration is worthwhile.

For each iteration, p solutions are calculated, creating p residuals and p new vectors. As long as these residuals have positive norms within the precision of the calculation, they should be added to the subspace, as discussed later in this chapter with a concrete example.

Thus each iteration has its most expensive step being the calculation of $n \times p$ products from $n \times p$ vectors. Calculating even one more solution at every iteration is a significant cost for large n.

1.2.4 Reference Implementation

As an example of a current, actively developed software for these equations, `escf` in TURBOMOLE[4] can be examined as an example of what hard coding a Krylov Subspace Algorithm. For flags specific to the response equations, there are 12 user-set flags for non-relativistic calculations (and 7 flags for relativistic options) that go through complicated and nested logic loops to ensure consistency and initialize other flags. The resulting flags change the behavior of the Krylov Subspace solver `respon` by calling different subroutines in `respon` depending on the particular flags. The number of solutions requested is also set separately by the user in a different variable.

None of the input for `escf` can be used to switch between the options listed for *libkrylov*, except to change the number of solutions requested. Adding new features typically involving adding another flag to the complicated loops, and implementing new subroutines since extending present subroutines without breaking present features is difficult.

An array, `diag`, is hard-coded to contain the zeroth order solutions, and passed into the Krylov subspace solver subroutine `respon` - `respon` is hard coded to use the `diag` for start vectors of the lowest eigenvalues, and may use a different array `sqdiag` for preconditioning, depending on the flags, where `sqdiag` is expected to be consistent with the element-wise square of `diag` array.

The steps of the solve have different, alternative, subroutines or code blocks depending on the flags. For example, there are 2 different subroutines for step 3,4 and 5 of the generic algorithm, (steps which are combined for efficiency,) but result in each subroutine having special features depending on the flags: `ritz` and `ritz2`. The case of complex solutions has 3 other subroutines: `gritz`, `gegdvd`, and `gritz2`. Even within TURBOMOLE, comparing between these subroutines and implementing new options consistently among them is difficult.

As the solve runs, only the residual norm and the number of solutions converged is reported to the user. There is only one error message obvious to the user: "`reference state unstable`".

1.2.5 *libkrylov* design decisions

Led by Rappoport, the latest design of *libkrylov* uses modern Fortran features for object-orientated programming. Expecting concurrent large problems that have matrix-vector projects that can be efficiently calculated together, `space` type objects are implemented with setters and getters to pass vectors to a matrix vector product matching the interface described in code snippet in figure 1.1.

The `space` type objects contain other objects relating to the type of matrix elements and the kind of problem, built in a modular fashion to be easily extended to include efficient memory management tools or special file input/output options. The feature list of `space` type objects

```
function krylov_i_real_multiply(full_dim, subset_dim, vectors, products) &
                           result(error)
  integer, intent(in) :: full_dim, subset_dim
  real, intent(in) :: vectors(full_dim, subset_dim)
  real, intent(out) :: products(full_dim, subset_dim)
  integer :: error
end function krylov_i_real_multiply

function krylov_i_real_block_multiply(num_spaces, total_size, full_dims, &
                                subset_dims, offsets, vectors, products) &
                                result(error)
  integer, intent(in) :: num_spaces, total_size, full_dims(num_spaces), &
                         subset_dims(num_spaces), offsets(num_spaces)
  real, intent(in) :: vectors(total_size)
  real, intent(out) :: products(total_size)
  integer :: error
end function krylov_i_real_block_multiply
```

Figure 1.1: Function interface for matrix vector product functions accepted by *libkrylov*

are still under active development and expected to grow under further development. The `space` type objects are initialized with data and selected options before calling the solve. This chapter focuses on all `space`s sharing a set of objects modified by user-selected options that affect the progression of the Krylov Subspace Algorithm - which allow for consistent behavior between spaces. There is a need to examine these options, in particular their interactions with each other, analytically and on a test problem.

1.3 Characterizing options for a Krylov Subspace Solve

A definition of the options available in and their expected importance to the behavior of the solve is presented, before these options are used on a test problem.

1.3.1 Starting vectors

With this design of *libkrylov*, initial vectors are set before calling the solver. (This step is within the Krylov subspace solver of `escf`.) The initial vectors $\mathbf{V}^{[1]}$ should be set to maximize their overlap with the desired solutions – the difference in the overlap of the initial vectors and the desired solutions is expected to change the convergence behavior of the solve.

There is the option to choose between $\mathbf{A}^{[0]}$ and $\mathbf{D}$; unit vectors based on the smallest (diagonal) elements of these (diagonal) matrices are probably a good estimate for the eigenvectors of the smallest eigenvalues. In the best-case scenario there is no difference in the resulting $\mathbf{V}^{[1]}$ between these two options.

$\mathbf{A}^{[0]}$ is the more efficient option. If using the same matrix-vector product engine to generate $\mathbf{D}$, the cost of calculating the full $\mathbf{D}$ would be the same as n matrix-vector products on the full space. This cost of calculating the full $\mathbf{D}$ may be expensive, weighted against the cost of additional matrix-vector products for the solve to converge (if it converges at all, after restarts). More efficient options for calculating a small segment of $\mathbf{D}$ (based on $\mathbf{A}^{[0]}$) may be possible. In particular, if the matrix-vector products are saved for a calculation that fails to converge, the first iteration's products of unit vectors allow for the calculation of a small segment of $\mathbf{D}$ (which is made straightforward in *libkrylov* via getter functions). For this chapter's test, $\mathbf{D}$ is stored on file and can be compared directly to $\mathbf{A}^{[0]}$.

More interestingly, using $\mathbf{D}$ is similar to the Epstein-Nesbet pair approximation (for orbital energies, applied to orbital energy gaps here).[17] It is known that including the 4 center integrals results in the diagonal values no longer being size-extensive, (unlike the eigenvalues,) as more electron hole pairs considered would increase the contribution of these integrals to $\mathbf{D}$. For maximizing overlap with the desired eigenvectors for a fixed number of starting vectors, unit vectors based on the $\mathbf{D}$ may be better as it accounts for some of the energy interactions between electron-hole-electron-hole pairs, and may be worse because not all of these interactions are considered. However, for increasing the number of starting vectors to maximize the overlap with the desired eigenvalues, the accounting of some of the energy interactions by $\mathbf{D}$ could more systematically improve the initial guess compared to $\mathbf{A}^{[0]}$. Repeating the above statement, the cost of calculating the full $\mathbf{D}$ may be too expensive.

1.3.2 Super matrix notation

To aid later analysis, it is useful to use the projection basis vectors $\mathbf{V}$, to define the vectors outside of the projection as

$$\mathbf{J}\mathbf{J}^{\dagger} = \mathbf{1} - \left(\mathbf{V}(\mathbf{V}^{\dagger}\mathbf{V})^{-1}\mathbf{V}^{\dagger}\right) \tag{1.20}$$

where normalization is required to ensure $\mathbf{J}$ is perpendicular to $\mathbf{V}$. For generality we do not assume $\mathbf{V}$ is orthonormalized. $\mathbf{J}$ are never explicitly computed as a full set, so for simplicity we will set $\mathbf{J}$ as an orthonormal set.

We can introduce a super-matrix notation:

$$\begin{pmatrix} \mathbf{V} \\ \mathbf{J} \end{pmatrix}^{\dagger} \begin{pmatrix} (\mathbf{V}^{\dagger}\mathbf{V})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{J} \end{pmatrix} = \mathbf{1} \tag{1.21}$$

And so write any (operator) matrix as:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & \mathbf{Z}_2 \\ \mathbf{Z}_3 & \mathbf{Z}_4 \end{pmatrix} \tag{1.22}$$

where component 1 projects within the iteration's subspace, component 2 projects from outside the subspace into the subspace, component 3 projects from inside the subspace to outside the subspace, and component 4 projects within the space outside the subspace.

Component 1 and 4 are well-described as "projected operators" where these components respectively act exclusively within the subspace and outside the subspace, rotating vectors within those spaces. Component 2 and 3 are more interesting in that they rotate vectors into and out of the subspace respectively. Note that all components usually scale the vectors in someway, and that most operators do not have $\mathbf{V}$ as eigenvectors.

To aid in the further discussion, note that:

$$\mathbf{AV} = \begin{pmatrix} \mathbf{A}_1\mathbf{V} \\ \mathbf{A}_3\mathbf{V} \end{pmatrix} \tag{1.23}$$

$$\mathbf{X}\boldsymbol{\omega} = \begin{pmatrix} \mathbf{X}\boldsymbol{\omega} \\ \mathbf{0} \end{pmatrix} \tag{1.24}$$

Having noted that solutions are within a subspace, it is worth pointing out that for different solutions on the full space to be comparable, (even solutions of problems differing by their RHS,) they need to be solved on the same subspace. This is another key reason for multiple matrix vector products to be evaluated at once, and for adding multiple new vectors to the subspace at each iteration.

1.3.3 Ideal new vectors and their orthonormality

Besides the choice of starting vector, two closely related options in the Krylov subspace solve relate to the treatment of new vectors: the orthornormalization scheme to conserve the eigenvalues of the projected coefficient matrix, and the preconditioning of residuals to generate new vectors. These options are both set by passing a corresponding character into setter functions.

As mentioned and discussed in the literature, the most straightforward selection of new vectors $\mathbf{T}$ is to use the raw residuals $\mathbf{R}$ In other words, $\mathbf{T}^{null} = \mathbf{R}$.

$$\mathbf{R} = \begin{pmatrix} \mathbf{0} \\ \mathbf{R} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \hat{\mathbf{A}}_3\mathbf{X} \end{pmatrix} \tag{1.25}$$

The residuals $\mathbf{R}$ are thus orthogonal to the projection basis $\mathbf{V}$ of the same iteration - however the columns of $\mathbf{R}$ can have linear dependence among themselves.

Sleijpen and Genseberger can be credited for defining ideal new vectors $\check{\mathbf{T}}$ such that

$$\mathbf{X} \approx \mathbf{X}^{[k]} + \check{\mathbf{T}}^{[k]} = \begin{pmatrix} \mathbf{X}^{[k]} \\ \check{\mathbf{T}}^{[k]} \end{pmatrix} \tag{1.26}$$

where the solution of the direct solve can be approximated by the solution of the current

Krylov iteration and the addition of a new vector $\check{\mathbf{T}}$ (barring normalization).[8, 9]

Substituting this above expression into the exact eigenvalue problem equation:

$$\begin{pmatrix} \acute{\mathbf{A}}_1 & \acute{\mathbf{A}}_2 \\ \acute{\mathbf{A}}_3 & \acute{\mathbf{A}}_4 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \check{\mathbf{T}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{1.27}$$

It can be determined that

$$\acute{\mathbf{A}}_2 \check{\mathbf{T}} = \mathbf{0} \tag{1.28}$$

and

$$\acute{\mathbf{A}}_3 \mathbf{X} + \acute{\mathbf{A}}_4 \check{\mathbf{T}} = \mathbf{0} \tag{1.29}$$

. (For the general problem, $\acute{\mathbf{A}}_2 \check{\mathbf{T}} = \mathbf{0}$ and $\acute{\mathbf{A}}_3 \mathbf{X} + \acute{\mathbf{A}}_4 \check{\mathbf{T}} = \mathbf{P}_{(J)}$, which lead to essentially the same conclusions.)

As a small side note, from equation 1.28, $\mathbf{V}^\dagger \acute{\mathbf{A}}_2 \check{\mathbf{T}} = ((\acute{\mathbf{A}}^{-1})_3 \mathbf{V})^\dagger \check{\mathbf{T}} = \mathbf{0}$, a condition that is difficult to fulfill with approximations to $\acute{\mathbf{A}}$.

As can be seen from the equation 1.29, $\hat{\mathbf{A}}_3 \mathbf{X} = \mathbf{R} = -\acute{\mathbf{A}}_4 \check{\mathbf{T}}$, and since in general $\acute{\mathbf{A}}_4 \neq \mathbf{1}$, $\mathbf{R} \neq c\check{\mathbf{T}}$ for a scalar c. Instead, $\check{\mathbf{T}} = -(\acute{\mathbf{A}}^{-1})_4 \mathbf{R}$. An interpretation is considering $\hat{\mathbf{A}}$ as a function, the residuals as output of that function, and $\check{\mathbf{T}}$ is what is desired to improve the input into the function. Convergence can be slow with the raw residual as new vectors, and at the same time $-(\acute{\mathbf{A}}^{-1})_4$ is inaccessible. The standard approach in the literature has been to precondition the residuals with a preconditioner $\acute{\mathbf{M}}$ that both approximates $\acute{\mathbf{A}}$ (specifically $\acute{\mathbf{A}}_4$) and is easily inverted, to obtain the new vectors as preconditioned residuals $\mathbf{T}$:

$$\acute{\mathbf{M}}\mathbf{T} = \mathbf{R}$$

Using either $\mathbf{A}^{[0]}$ or $\mathbf{D}$ as $\mathbf{M}$ are choices available to users of *libkrylov*, by the user choosing to pass either matrix to the same preconditioner object of `space`. A generic $\acute{\mathbf{M}}$ is used when discussing the preconditioner and its implementation. If the specific choice of $\mathbf{A}^{[0]}$ or $\mathbf{D}$ is important, $\acute{\mathbf{M}}^{[0]}$ and $\acute{\mathbf{M}}^{D}$ can be used.

The preconditioned residuals (or new vectors), $\mathbf{T}$, are typically no longer orthogonal to the subspace of that iteration (due to the approximate nature of $\acute{\mathbf{M}}$), in addition to retaining linear dependence of among its columns. This linear dependence is treated differently by different orthonormalization schemes, convoluting the effects of the choice of preconditioner and the choice of orthonormalization scheme. The different orthonormalizers and their treatment of linear dependence is discussed first to better characterize the behavior of preconditioners with these different orthonormalization schemes.

1.3.4 Orthonormalization schemes

The linear dependence of the new vectors among themselves and with the projection basis vectors of that iteration must be tackled in the algorithm. The consequence of either linear dependence is the same – the Krylov subspace problem no longer shares properties of the full space problem. [18] Many approximate iterative solvers encounter numerical instability due to linear dependence, with restarts of the algorithm as the consistent solution. [18] However, restarts may prevent convergence, thus reducing the build up of linear dependence is still crucial. [18]

The straightforward approach is to ensure that all projection basis vectors $\mathbf{V}$ are orthonormal (to numerical precision) at every step by orthogonalizing and normalizing the new vectors against the projection basis vectors of the iteration. This costly procedure is treated as a default in the literature about various preconditioners.

Keeping an Orthonormal Krylov projection basis

A typical orthonormalization scheme that is used this chapter is the modified Gram-Schmidt procedure with normalization in Figure 1.2 as described by Saad.[15] This is not the typical forward looking modified Gram-Schmidt procedure described elsewhere in the literature.[18]

```
do l = 1_IK, n_dim
    do j = 1_IK, l - 1_IK
        inner_product = ddot(m_dim, array(:, j), 1_IK, array(:, l), 1_IK)
        array(:, l) = array(:, l) - ( inner_product * array(:, j) )
    end do
    normsq = ddot(m_dim, array(:, l), 1_IK, array(:, l), 1_IK)
    array(:, l) = array(:, l) / sqrt(normsq)
end do
```

Figure 1.2: Saad's modified Gram Schmidt with normalization

Keeping an Non-orthonormal Krylov projection basis

The alternative proposed by Furche in 2016[3] is to keep the new vectors with their decreasing norms and instead construct a gram matrix $\mathbf{s}$(typically referred to as an overlap matrix in chemistry). The diagonals of $\mathbf{s}$, $\mathbf{d}$, can be used to scale the Gram matrix to improve the numerical precision. The scaled Gram matrix, $\mathbf{d}^{-0.5}\mathbf{s}\mathbf{d}^{-0.5}$, has unit diagonals.

$$\mathbf{s} = \mathbf{V}^{\dagger}\mathbf{V} \tag{1.30}$$

$$\mathbf{d} = \mathrm{diag}(\mathbf{s}) \tag{1.31}$$

After scaling the Gram matrix for numerical stability, a Cholesky decomposition can be carried out.

$$\mathbf{L}\mathbf{L}^{\dagger} = \mathbf{d}^{-0.5}\mathbf{s}\mathbf{d}^{-0.5} \tag{1.32}$$

$$\bar{\mathbf{a}} = \mathbf{L}^{-1}\mathbf{d}^{-0.5}\mathbf{a}\mathbf{d}^{-0.5}\mathbf{L}^{-\dagger} \tag{1.33}$$

The Cholesky decomposition and diagonals of the Gram matrix can be used to transform the projected Krylov subspace problem into an orthonormal basis as described in equation 1.33. This is computational cheaper compared to orthonormalizing new vectors, as most of the operations are now carried out on the subspace instead of the full space. This also allows the matrix-vector product function to screen out more elements of the new vectors as their magnitudes are no longer inflated by normalization - preconditioned residuals typically have reducing norms as the iterations proceed. Such screening is expected to affect the convergence behavior of the solver and has been reported in the literature for specific problems,[3] but should not be relied on in a solver meant for more general cases. The linear dependence of the new vectors among themselves and with the previous projection basis vectors are tracked together in the condition number of the scaled Gram matrix $\mathbf{d}^{-0.5}\mathbf{s}\mathbf{d}^{-0.5}$. When the inverse condition number, calculated with the Forbenius norm of the scaled Gram matrix, decreases below a threshold, the Cholesky decomposition and transformation of the problem are taken to be no longer numerically stable and the iterations are stopped.

As the linear dependence of the projection vectors on the full space is no longer eliminated, but carried through the iterations, distinguishing the two sources of linear dependence - of the new vectors among themselves and of the new vectors with the previous projection basis vectors, becomes important, in particular if the preconditioner causes preconditioned residuals to have significant linear dependence among themselves or significant components in the subspace, respectively. The preconditioners are described below, distinguishing between sources of linear dependence.

A Non-orthonormal Krylov projection basis with less linear dependence

A new feature reported in this chapter is a middle ground between orthonormalization of all projection basis vectors and the non-orthonormal Krylov subspace proposed by Furche. The linear dependence of the new vectors among themselves can be eliminated by SVD of the new vectors among themselves, with further efficiency improvements by carrying out a QR decomposition first. Doing so does not orthogonalize the new vectors against the present subspace, but improves the condition of the new basis if the preconditioner projects the residuals onto the same vector.

$$\mathbf{Q}\mathbf{R}' = \mathbf{T} \tag{1.34}$$

$$\mathbf{U}\Sigma\mathbf{W}^\dagger = \mathbf{R}' \tag{1.35}$$

$$\tilde{\mathbf{T}} = \mathbf{Q}\mathbf{U}\Sigma \tag{1.36}$$

Independent option of Projection Basis scheme

With these different orthonormalization schemes, the significance of the linear dependence of the preconditioned residuals over the iterations is different between them, hence it is reasonable to have the preconditioner and orthonormalization schemes as separate options that can be toggled independently. The orthonormalization steps of the algorithm are implemented in an orthonormalizer object in `space`.

1.3.5 Preconditioning of residuals

In light of the non-orthonormal Krylov subspace algorithm, the preconditioning present in the literature are examined for their contribution to the linear dependence of the full projection basis.

Davidson Preconditioning

Davidson suggested that the diagonal of $\mathbf{A}$ ($\mathbf{D}$) as a good preconditioner, being a good approximation for $\mathbf{A}$, and is easily inverted with eigenvalues from the subspace solve.[19] As discussed, the diagonals of $\mathbf{A}$ are usually not calculated explicitly, the diagonals of $\mathbf{A}^{[0]}$ are used instead:

$$\acute{\mathbf{M}}\mathbf{T} = \mathbf{M}\mathbf{T} - \mathbf{T}\boldsymbol{\Omega} \tag{1.37}$$

$$\acute{\mathbf{M}}\mathbf{T}^{Davidson} = -\mathbf{R} \tag{1.38}$$

Conjugate gradient preconditioning is an option to consider,

$$\mathbf{M}\mathbf{T}^{ConjugateGradient} = -\mathbf{R}$$

but steps away from equation 1.29. Conjugate gradient preconditioning is implemented in *libkrylov*.

It has been noted by Davidson, Sleijpen and Genseberger that the Davidson preconditioning using $\mathbf{D}$ becomes unstable near convergence since $D_{kk} \approx \Omega_k$ near convergence of that vector k, for diagonally dominant matrices. [19, 8, 9] While *libkrylov* will have regularization to avoid such instability, this chapter's test will exit the solver with an error if this instability is detected.

To compare to equation 1.29, in our notation,

$$-\acute{\mathbf{M}}^{-1}\mathbf{R} = \begin{pmatrix} -(\acute{\mathbf{M}}^{-1})_2\mathbf{R} \\ -(\acute{\mathbf{M}}^{-1})_4\mathbf{R} \end{pmatrix} = \begin{pmatrix} -(\acute{\mathbf{M}}_3)^{-1}\mathbf{R} \\ -(\acute{\mathbf{M}}_4)^{-1}\mathbf{R} \end{pmatrix} \tag{1.39}$$

ignoring the small finite precision error that would result in contributions from other components of $\mathbf{M}^{-1}$. Our notation clearly shows that $-(\acute{\mathbf{M}}^{-1})_2\mathbf{R}$ is present in the preconditioned residual, which is linearly dependent on the Krylov subspace projection basis vectors of the same iteration.

To make some subsequent expansions clearer:

$$\mathbf{T} = -\acute{\mathbf{M}}^{-1}\mathbf{R} = \begin{pmatrix} -(\acute{\mathbf{M}}^{-1})_2\hat{\mathbf{A}}_3\mathbf{X} \\ -(\acute{\mathbf{M}}^{-1})_4\hat{\mathbf{A}}_3\mathbf{X} \end{pmatrix} = \begin{pmatrix} -(\acute{\mathbf{M}}_3)^{-1}\hat{\mathbf{A}}_3\mathbf{X} \\ -(\acute{\mathbf{M}}_4)^{-1}\hat{\mathbf{A}}_3\mathbf{X} \end{pmatrix} \tag{1.40}$$

Genseberger has shown that the space formed by these preconditioning methods does form a different Krylov subspace, $\mathcal{K}_k(\acute{\mathbf{M}}^{-1}\hat{\mathbf{A}}, \mathbf{V}^{[1]})$. Questions about the properties of the "full space" $\mathcal{K}_n(\acute{\mathbf{M}}^{-1}\hat{\mathbf{A}}, \mathbf{V}^{[1]})$ are not investigated in this chapter.

A concern raised by Sleijpen and Genseberger about Davidson preconditioning is that $(\acute{\mathbf{M}}^{-1})_2\mathbf{R} \approx \mathbf{V}\mathbf{V}^{\dagger}\mathbf{X} = \mathbf{X}$ can dominate the resulting preconditioned residuals, resulting in lost precision and slower convergence even after orthogonalization; specifically, that orthogonalization of the preconditioned residuals will not be able to recover good new vectors. [8, 9] For an orthonormal Krylov Subspace algorithm, this would results in numerically zero new vectors after orthornormalization.

Extending the above to a non-orthonormal Krylov projection basis, this linear dependence would be one contribution to depreciation of the condition number of the scaled Gram matrix - leading to numerical instability in the subspace. This difference in concerns between the orthonormal Krylov subspace projection and non-orthonormal Krylov subspace projection has yet to be reported in the literature. While numerically zero new vectors of the orthonormal Krylov subspace projection can be resolved by adding terms to the preconditioning to increase the magnitude of the preconditioned residuals outside the present subspace, an unstable orthonormalization of the subspace in the non-orthonormal Krylov subspace projection requires reducing the linear dependence of the preconditioned residuals among themselves and with the present subspace projection basis.

It is useful to note that the projection of residuals onto the solutions by Davidson Preconditioning is more significant the better $\mathbf{M}$ is as an approximation of $\mathbf{A}$, particularly close to convergence when the useful component outside the subspace becomes small (and where the eigenvalues of the subspace solve also approach the eigenvalues of the direct solve). *A better approximation for $\mathbf{A}^{-1}$ in Davidson preconditioning can be more unstable with a non-orthonormal Krylov subspace projection basis.* Attempts to avoid this include work by Huang where the problem is partitioned with part of the problem solved directly.[20]

Another means to avoid the instability from projection of the residuals onto their solutions is the use of Jacobi-Davidson Preconditioners as defined by Sleijpen and Genseberger with an orthogonal Krylov subspace method - where the preconditioner matrix has components in the projection basis projected out.

First Variant of Sleijpen's Jacobi-Davidson Preconditioner

Sleijpen's and Genseberger's attempts use projectors such as $(\mathbf{1} - \mathbf{X}\mathbf{X}^{\dagger})$ to fulfil these conditions. [7, 8, 9] Sleijpen writes his preconditioner for some generic $\acute{\mathbf{M}}$, so the choice of $\mathbf{A}^{[0]}$ or $\mathbf{D}$ to define $\acute{\mathbf{M}}$ are both valid and is a way to compare $\mathbf{A}^{[0]}$ and $\mathbf{D}$. It is also possible to not use the eigenvalues of the subspace solve and use just $\mathbf{M}$ (similar to conjugate gradient preconditioning), but that is not yet implemented.

Sleijpen defines his preconditioner (for normalized solutions) :

$$(\mathbf{1} - \mathbf{X}\mathbf{X}^{\dagger})\acute{\mathbf{M}}(\mathbf{1} - \mathbf{X}\mathbf{X}^{\dagger})\mathbf{T} = -\mathbf{R} \tag{1.41}$$

We can add in additional normalization, since Sleijpen assumed orthonormal solutions and vectors.

$$(\mathbf{1} - (\mathbf{X}(\mathbf{X}^{\dagger}\mathbf{X})^{-1}\mathbf{X}^{\dagger}))\acute{\mathbf{M}}(\mathbf{1} - (\mathbf{X}(\mathbf{X}^{\dagger}\mathbf{X})^{-1}\mathbf{X}^{\dagger}))\mathbf{T} = \mathbf{R} \tag{1.42}$$

Sleijpen obtains his one step formula by expanding one of the projectors. The subsequent approximation breaks the requirement that $\mathbf{T}$ to be orthogonal to $\mathbf{V}$, and can be seen as an interpolation of Jacobi "shift and invert" approaches and Davidson preconditioning. Assuming the focus is on only the corresponding solution vector for each residual, the first variant of Sleijpen's Jacobi-Davidson preconditioning is described below. This first variant of Sleijpen's Jacobi-Davidson preconditioning is focused on in this chapter, with the other variants of Jacobi-Davidson described to contrast with this first variant.

$$\mathbf{T}_j^{Sleijpen,1st} = \frac{\mathbf{X}_j^\dagger \acute{\mathbf{M}}^{-1}\mathbf{R}_j}{\mathbf{X}_j^\dagger \acute{\mathbf{M}}^{-1}\mathbf{X}_j}\acute{\mathbf{M}}^{-1}\mathbf{X}_j - \acute{\mathbf{M}}^{-1}\mathbf{R}_j$$

The super-matrix notation is useful to separate the different components of the matrices for Sleijpen's first variant:

$$\mathbf{T}_j = \begin{pmatrix} \frac{\mathbf{X}_j^\dagger (\acute{\mathbf{M}}^{-1})_2\mathbf{R}_j}{\mathbf{X}_j^\dagger (\acute{\mathbf{M}}^{-1})_1\mathbf{X}_j}(\acute{\mathbf{M}}^{-1})_1\mathbf{X}_j - (\acute{\mathbf{M}}^{-1})_2\mathbf{R}_j \\ \frac{\mathbf{X}_j^\dagger (\acute{\mathbf{M}}^{-1})_2\mathbf{R}_j}{\mathbf{X}_j^\dagger (\acute{\mathbf{M}}^{-1})_1\mathbf{X}_j}(\acute{\mathbf{M}}^{-1})_3\mathbf{X}_j - (\acute{\mathbf{M}}^{-1})_4\mathbf{R}_j \end{pmatrix} \tag{1.43}$$

The term on the subspace $\mathbf{T}_{j(V)}$ is orthogonal to the corresponding solution for Sleijpen's first variant - which can be seen by multiplying the solution vector $\mathbf{X}_j^\dagger$ from the left. In the case that the preconditioner projects the residual back onto its corresponding solution, the first variant of Sleijpen is useful to cancel out that projection. However, the term on the subspace $\mathbf{T}_{j(V)}$ is not orthogonal to other solutions - while $\mathbf{X}^\dagger\mathbf{X} = \mathbf{1}$, typically $\mathbf{X}^\dagger\acute{\mathbf{M}}^{-1}\mathbf{X}$ is not diagonal unless there are special features of $\acute{\mathbf{M}}$. The trade-off in eliminating linear dependence with one solution vector while increasing linear dependence with other vectors in the subspace can be evaluated by following the condition number. This linear dependence of the preconditioned residuals with the projection basis vectors is only eliminated by orthogonalizing the new vectors against the projection basis vectors.

How the subspace is expanded on the next iteration by $\mathbf{T}_{(J)}$, the components outside the subspace, is also important. The term outside the subspace now has an additional component $(\acute{\mathbf{M}}^{-1})_3\mathbf{X}_j$ - whose magnitude increases if the preconditioner matrix $\acute{\mathbf{M}}^{-1}$ projects the residuals onto that solutions vector, and decreases if the preconditioner projects that solution onto itself. There is no consideration of increased linear dependence in $\mathbf{T}$ due to that term. This makes sense when considering that these preconditioners were written for orthonormal Krylov Subspace projection bases – and as this linear dependence is among the new vectors, the SVD of new vectors will also remove this linear dependence.

The other variants of Jacobi-Davidson are described to contrast the first variant of Sleijpen preconditioning.

Other Variants of Jacobi-Davidson Preconditioner

Following from the initial definition of the projector over all solutions, it is possible to write a second variant of Sleijpen preconditioning:

$$\begin{aligned}\mathbf{T}_j^{Sleijpen,2nd} &= \sum_k \left(\frac{\mathbf{X}_k^\dagger \acute{\mathbf{M}}^{-1} \mathbf{R}_j}{\mathbf{X}_k^\dagger \acute{\mathbf{M}}^{-1} \mathbf{X}_k} \acute{\mathbf{M}}^{-1} \mathbf{X}_k \right) - \acute{\mathbf{M}}^{-1} \mathbf{R}_j \\ &= \sum_k \left(\frac{\mathbf{X}_k^\dagger (\acute{\mathbf{M}}^{-1})_2 \mathbf{R}_j}{\mathbf{X}_k^\dagger (\acute{\mathbf{M}}^{-1})_1 \mathbf{X}_k} \left((\acute{\mathbf{M}}^{-1})_3 \mathbf{X}_k + (\acute{\mathbf{M}}^{-1})_1 \mathbf{X}_k \right) \right) - (\acute{\mathbf{M}}^{-1})_2 \mathbf{R}_j - (\acute{\mathbf{M}}^{-1})_4 \mathbf{R}_j \qquad (1.44)\end{aligned}$$

Where the overlap of all preconditioned solutions with each residual is considered. Thus, the magnitude of the preconditioned residual outside the Krylov Subspace, $(\acute{\mathbf{M}}^{-1})_3 \mathbf{X}_k$, are greater – if the preconditioner projects solutions onto residuals but not onto itself. This also results in increased linear dependence of the preconditioned residuals among themselves, another potential problem taken care of by an orthonormal Krylov projection basis or an SVD of preconditioned residuals. Within the subspace, the summation also results in the preconditioned residuals from this second variant not being orthogonal to any solution vectors of that iteration. With an orthonormal Krylov subspace projection, the second variant of Sleijpen would help project new vectors out of the subspace, and is not implemented in *libkrylov*.

Genseberger notes that projecting out only the solution vectors still leaves some components of $\mathbf{T}$ in the subspace, but reasons that orthogonalization by an orthonormal projection basis method will remove them.

Genseberger extends the idea of projecting out solution vectors from the preconditioner to projecting out basis vectors in a subsequent paper[9]:

$$(\mathbf{1} - \mathbf{V}\mathbf{V}^{\dagger})\acute{\mathbf{M}}(\mathbf{1} - \mathbf{V}\mathbf{V}^{\dagger})\mathbf{T} = -\mathbf{R} \tag{1.45}$$

We can also add in additional normalization to the above equation, since Genseberger also assumed orthonormal projection basis vectors.

$$(\mathbf{1} - (\mathbf{V}(\mathbf{V}^{\dagger}\mathbf{V})^{-1}\mathbf{V}^{\dagger}))\acute{\mathbf{M}}(\mathbf{1} - (\mathbf{V}(\mathbf{V}^{\dagger}\mathbf{V})^{-1}\mathbf{V}^{\dagger}))\mathbf{T} = \mathbf{R} \tag{1.46}$$

Genseberger does not state a form for his preconditioner, but it is straightforward to extend Sleijpen's formalism to obtain the first variant of Genseberger:

$$\mathbf{T}_j^{Genseberger,1st} = \sum_k \left(\frac{\mathbf{V}_k^{\dagger}\acute{\mathbf{M}}^{-1}\mathbf{R}_j}{\mathbf{V}_k^{\dagger}\acute{\mathbf{M}}^{-1}\mathbf{V}_k} \acute{\mathbf{M}}^{-1}\mathbf{V}_k \right) - \acute{\mathbf{M}}^{-1}\mathbf{R}_j \tag{1.47}$$

There is no appropriate way to pick a single basis vector $\mathbf{V}_k$ that should be orthogonal to $\mathbf{T}_j$, so all vectors are considered and hence $\mathbf{T}_j$ is not orthogonal to any basis vector, comparable to Sleijpen's second variant. Also comparable to Sleijpen's preconditioning, there are now $(\acute{\mathbf{M}}^{-1})_3\mathbf{V}$ components added to the subspace, which may improve the convergence.

It is worth pointing out that using an orthonormal projection basis does not ensure that $\mathbf{V}_j^\dagger \acute{\mathbf{M}}^{-1}\mathbf{V}_k = 0$ for $j \neq k$ – Genseberger's first variant is not an orthonormalization scheme for the preconditioned residual, and depend heavily on the properties of $\acute{\mathbf{M}}$. Genseberger's first variant is not implemented in *libkrylov*.

Given the expense of matrix-vector and vector-vector products within larger Krylov subspaces, and the lack of orthogonality properties of the second variant of Sleijpen and the first variant of Genseberger, the first variant of Sleijpen is both cheaper and more useful than the other Jacobi-Davidson presented so far.

This also suggests that it would be cheaper to combine all solution vectors or basis vectors into a third variant of Sleijpen or second variant of Genseberger:

$$\mathbf{X}^{summed} = \sum_k^p \mathbf{X}_k$$

$$\mathbf{T}_j^{Sleijpen,3rd} = \frac{\mathbf{X}^{summed,\dagger}\acute{\mathbf{M}}^{-1}\mathbf{R}_j}{\mathbf{X}^{summed,\dagger}\acute{\mathbf{M}}^{-1}\mathbf{X}^{summed}}\acute{\mathbf{M}}^{-1}\mathbf{X}^{summed} - \acute{\mathbf{M}}^{-1}\mathbf{R}_j$$

$$\mathbf{V}^{summed} = \sum_k^q \mathbf{V}_k$$

$$\mathbf{T}_j^{Genseberger,2nd} = \frac{\mathbf{V}^{summed,\dagger}\acute{\mathbf{M}}^{-1}\mathbf{R}_j}{\mathbf{V}^{summed,\dagger}\acute{\mathbf{M}}^{-1}\mathbf{V}^{summed}}\acute{\mathbf{M}}^{-1}\mathbf{V}^{summed} - \acute{\mathbf{M}}^{-1}\mathbf{R}_j$$

Given that $\mathbf{X}^\dagger\acute{\mathbf{M}}^{-1}\mathbf{X}$ and $\mathbf{V}^\dagger\acute{\mathbf{M}}^{-1}\mathbf{V}$ are not diagonal, these last variants of Jacobi Davidson have preconditioned residuals of greater magnitude inside and outside the subspace. In the case that convergence is slow but the condition number of the basis is relatively good, that greater magnitude may improve convergence without catastrophically destroying the condition of the basis, more efficiently than the second variant of Sleijpen or the first variant of Genseberger. In particular for the second variant of Genseberger, if the summed vector $\mathbf{V}^{summed}$ is saved, only the new vectors need to be added to the sum. These are quite different from the literature, and are not implemented in *libkrylov*.

1.3.6 Combining orthonormalization and preconditioning

A broad outcome of this chapter is that *libkrylov* presents a way to switch easily between orthonormalization and preconditioning options independently for a Krylov subspace solve, including between $\mathbf{D}$ and $\mathbf{A}^{[0]}$ for preconditioning.

Between the options for SVD in orthonormalization and the first variant of Sleijpen preconditioning, it is now possible to vary the extent that linear dependence within the projection basis vectors is eliminated. Using only SVD of the new vectors does not remove the linear dependence of the new vectors on the projection basis vectors of that iteration. Using only Sleijpen's first variant removes the linear dependence of one solution vector (which usually has component in all projection basis vectors) from one preconditioned residual vector. Using Sleijpen's first variant with SVD of the new vectors reduces the magnitude of the solution vectors entering the SVD, which may also improve the quality of the new vectors after SVD. The condition number of the scaled Gram matrix presents an opportunity to compare the magnitude of these linear dependencies for specific preconditioners for specific problems.

List of options

To summarise all the options available in this implementation of *libkrylov*:

- Choice of preconditioning scheme - null, conjugate gradient, Davidson, Sleijpen variant 1.

- Choice of $\mathbf{M}$ used in preconditioning - either $\mathbf{D}$ or $\mathbf{A}^{[0]}$
- Choice of using $\mathbf{D}$ or $\mathbf{A}^{[0]}$ to generate starting vectors based on the elements with the lowest magnitude for a lowest eigenvalue problem solver.
- Choice of orthonormalization scheme - keeping an orthonormal Krylov subspace basis, keeping a non-orthonormal Krylov subspace basis, keeping a non-orthonormal Krylov subspace basis with SVD of new vectors.
- Number of solutions to solve for.
- Number of starting vectors.

Default options selected in Turbomole

The standard approach in Turbomole on a fresh start of the algorithm is to create starting vectors corresponding to the smallest values of $\mathbf{A}^{[0]}$, and determine the number of starting vectors as the smaller of two numbers - either the number of solutions multiplied by 2, or the number of solutions plus 8: $q^{[1]} = \min(2p, p + 8)$.

The other options that have been fixed in Turbomole are the choice of orthonormalization and the choice of matrix for preconditioning. Turbomole uses the non-orthonormal Krylov subspace approach for non-relativistic problems, and uses the zeroth order $\mathbf{A}^{[0]}$ for preconditioning.

1.4 Running *libkrylov* on a test system

1.4.1 Test example system

The nucleobases are conjugated organic molecules whose electronic excitations are of interest since their excitations are linked to the photo degradation of DNA.[21]

Uracil has its ground state geometry optimized with Kohn Sham density functional theory with the PBE GGA functional[22] and def2-SVP basis set.[23] The scf convergence threshold is set to 10^{-12}, the reduced one electron density convergence is set to 10^{-11}. The geometry convergence threshold was set to 10^{-6}. The electronic structure was converged again with the hybrid GGA PBE0,[24, 25] with the same parameters. By adjusting the source code of Turbomole, the full $\mathbf{A}$ and $\mathbf{A}^{[0]}$ with TDA are be printed as JSON dictionaries.

Because the matrices are printed out on file, the presented results do not depend on screening by a matrix-vector product function.

Information about the problem

Uracil has a full space n of $2987 = 29 \times 103$, from 29 doubly occupied orbitals (58 electrons) and 103 empty orbitals (total electronic basis size is $29 + 103 = 132$) which leads to the full matrix and diagonal filling out a 1.9GB and 240KB JSON file respectively.

Focusing on the lowest two eigenvectors of the linear response equation in the zero-field limit, and the direct solve gives the lowest eigenvector significant contribution from mainly the HOMO-1 to LUMO transition, with contributions from HOMO-3 to LUMO, HOMO-1 to LUMO+1, HOMO-2 to LUMO+3, and HOMO-1 to LUMO+5. The second lowest eigenvector has significant contribution from mainly the HOMO to LUMO transition, with contributions from HOMO-2 to LUMO, HOMO-2 to LUMO+1, HOMO-4 to LUMO, and

HOMO to LUMO+5.

The (inverse) diagonal dominance factor

$$\left(\sum_j |A_{kj}|\right) / |A_{kk}|$$

for all k is plotted in figure 1.3, showing that there are a few elements that make the problem not diagonally dominant, using the definition of diagonal dominance in the literature.

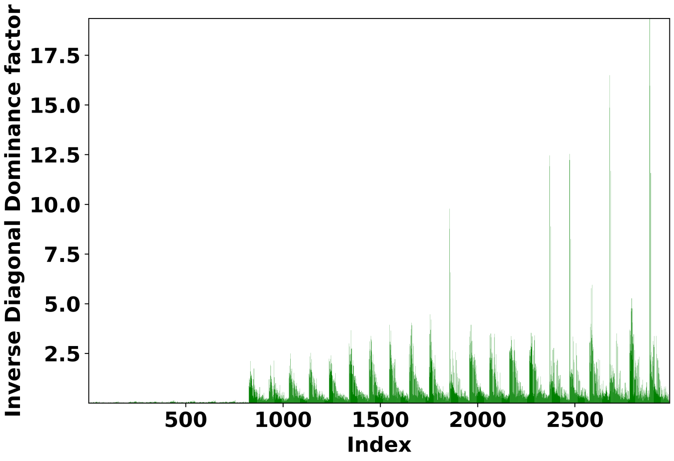

Figure 1.3: Inverse diagonal dominance factor of $\mathbf{A}$ for Uracil with PBE0

To distinguish between $\mathbf{A}$ and $\mathbf{A}^{[0]}$, the weighted magnitude of the two electron-hole terms

$$\left(A_{kk} - A^{[0]}_{kk}\right) / A^{[0]}_{kk}$$

for all k is plotted in figure 1.4, giving some indication of how close an approximation $\mathbf{A}^{[0]}$ is for $\mathbf{A}$. As can be noted, most of these contributions are negative (unlike with PBE, where they are positive). $\mathbf{A}^{[0]}$ is reasonably close to $\mathbf{D}$ for Uracil.

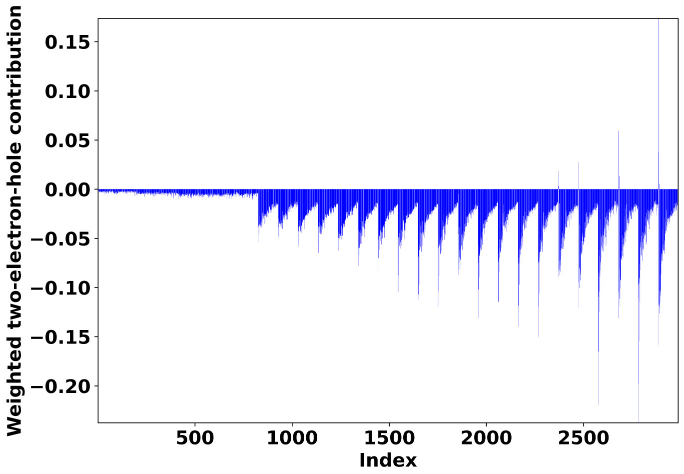

Figure 1.4: Comparing $\mathbf{A}^{[0]}$ and $\mathbf{D}$ for Uracil with PBE0

1.4.2 Presentation

For comparing residual norms with different options for the Krylov subspace solve method, the residual norms are numerically different, but the magnitude of the residual norms are more meaningful. By plotting the base 10 log of the residual norm, the similarity of similar methods is shown. The same applies to the inverse condition number. For example, switching between a non-orthonormal Krylov subspace method and a orthonormal Krylov subspace method, which does not change the component of the new vectors that improves the subspace projection, is shown in figure 1.5, for Uracil as described above, starting with 4 start vectors, solving for 2 solutions, using $\mathbf{A}^{[0]}$ for preconditioning and determining start vectors, with Davidson Preconditioning. It is worth nothing that the number of basis vectors at the end of these iterations is different (46 with orthonormal, 48 with non-orthonormal) despite the same behavior - a smaller number of basis vectors does not indicate a better solve, despite likely being computational cheaper.

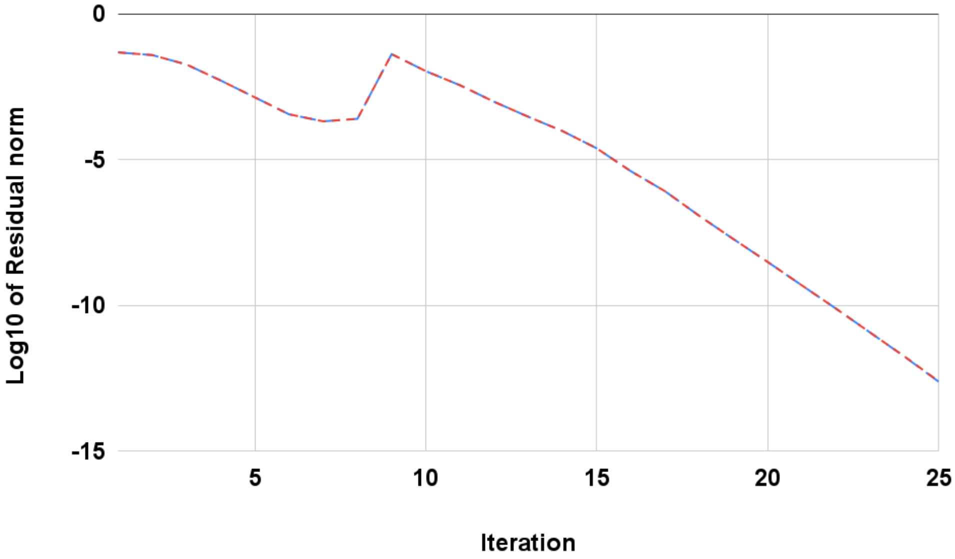

Figure 1.5: Comparing Orthonormal(red) versus Non-orthonormal(blue) Krylov Subspace basis for solving Uracil with PBE0, using $\mathbf{A}^{[0]}$ for preconditioning and determining start vectors, for 2 solutions with 4 starting vectors, Davidson preconditioning.

1.4.3 Choice of starting vectors

Ordering the Kohn-Sham occupied-virtual orbital energy differences by magnitude, the HOMO-LUMO energy gap is the smallest, followed by the HOMO-1 to LUMO energy gap. Using $\mathbf{A}^{[0]}$ to determine starting vectors, there are significant components of both lowest eigenvectors in the starting Krylov subspace basis. However, as seen by the discontinuity in figure 1.5, the convergence does not go smoothly. There is a “root-flip” when expanding the subspace sufficiently reorders the eigenvalues on the subspace, which can be seen by the magnitude of the basis vectors in the solutions as the solve progresses, as reported in table 1.1. In particular, the magnitude of basis vector 18, which were generated in iteration 7, are significant in the converged solution at iteration 25. The new vectors are in sequence of the corresponding solution and residual – for example the 17th basis vector is generated from the 1st solution vector of the previous iteration, and the 18th basis vector is generated from the 2nd solution vector of that same iteration.

Index	8	25
1	$2.6599491975435353E-003$	-0.88935141174254906
2	-0.20843773741198807	$2.8938744547479341E-006$
3	$1.5673774312419443E-003$	-0.27869094619054674
4	-0.79375180091822262	$-4.6302638986766570E-006$
5	-0.27789045851708838	$1.2745810735063956E-006$
6	0.45901364092333113	$4.7758313511509277E-006$
7	-0.10647163326569048	$-6.2147460762986442E-006$
8	-0.15196151610906061	$4.4784111994249937E-005$
9	$4.9783101228749987E-002$	$-3.0445816960467362E-005$
10	$-3.8417420338272112E-002$	$-4.0026268889248174E-005$
11	$9.3428002964870815E-003$	$5.3617830073128081E-004$
12	$-9.9989095674961007E-003$	$4.0262026340006449E-003$
13	$-1.6186359085475892E-003$	$3.0494348878124613E-003$
14	$-3.1957716521726761E-003$	$-6.6290801286772389E-003$
15	$2.4321735526475042E-004$	$3.2073203575574132E-002$
16	$1.4018619043448400E-003$	$-7.8897578957894635E-002$
17	$2.7042093487147754E-004$	$-5.3077305347891532E-003$
18	$-8.1467908234747813E-004$	0.13687955102785254

Table 1.1: Contributions of the first 18 basis vectors to the second lowest solution at iteration 8 and 25, from figure 1.5, for orthonormal Krylov subspace basis.
Horizontal lines are used to separate vectors from different iterations

Note that “root-flips” only occur for eigenvalue problems, not linear or shifted linear equations.

Using matrix **D**, the diagonals of **A**, the smallest element corresponds to the HOMO-1 to LUMO transition, and the HOMO to LUMO transition is the 7th smallest element. For most of the tests in this chapter, the number of starting vectors, $q^{[1]}$, if determined by matrix **D**, will be at least 7 to include significant overlap of the starting vectors with the desired solution.

As a reasonable example, the base 10 log of residual norm of the Krylov subspace solve for Uracil, using **D** for starting vectors and preconditioning, and with an orthornormal projection, is shown in figure 1.6. In solid blue is the line for 8 starting vectors. There are “root-flips” with most of the Krylov subspace solves – the Krylov subspace needs to increase

Index	6	22
1	0.20844514329562316	$-2.8938744410956401E-006$
2	0.79375568750700909	$4.6302639250764844E-006$
3	-0.48397471454934138	$-1.9609146909358363E-007$
4	-0.28335852691289559	$1.6398756300159254E-006$
5	$-1.3247007886289162E-004$	0.27869094619035650
6	$7.9693928276881598E-003$	$5.6654840669039743E-007$
7	$-1.7514923850620986E-004$	0.88935141174266574
8	$-1.4012205715174285E-005$	$-3.2645428643221244E-002$
9	$3.4605339723196217E-002$	$-1.4625250049507482E-006$
10	0.10278493351853409	$3.6375121882700367E-006$
11	$7.2111639958286422E-003$	$-1.5942278813202030E-006$
12	$9.7694758423946797E-003$	$-1.2102317658958906E-004$
13	$1.6764414636069541E-003$	$-8.4872855718686052E-005$
14	$1.6756154564802127E-003$	$-6.5653368321827888E-004$
15	$1.3116444102538957E-004$	$6.4188611902020736E-004$
16	$1.9753759499743681E-004$	$-2.2896551449373262E-002$
17	$2.4652087143832742E-005$	$-4.0405608226082171E-003$
18	$3.7804293083538815E-005$	-0.16851611886935974

Table 1.2: Contributions of the first 18 basis vectors to the second lowest solutions at iteration 6 and 22 starting with 8 start vectors, for an orthonormal Krylov subspace projection basis. This is the solid blue line from figure 1.6, Horizontal lines are used to separate vectors from different iterations

pass a certain size to find the correct second lowest eigenvector.

Comparing residual norm to error in eigenvalues

For the example, for the solid blue line in figure 1.6, the residual norm at iteration 6 is 1.60528 $\times 10^{-5}$, which indicates that the error in the eigenvalues is of that magnitude. However, the eigenvalues of that iteration are 0.17490723027469660 and 0.21032801489464834, which can be compared to the eigenvalues of a direct solve which are 0.174907230274533 and 0.205232307979602 - the second eigenvalue changes by 0.005095706915, which is two orders of magnitude larger than the residual norm.

The second lowest eigenvalue at iteration 6 turns out to correspond to the third lowest eigenvalue, which in the direct solve has an eigenvalue of 0.218950725008987.

For most applications, the exact ordering of eigenvectors is not very important, as long as the significant eigenvectors are captured. For non-adiabatic molecular dynamics that define adiabatic surfaces based on adiabatic excitation energy ordering, the relative order becomes more critical. In these applications, it is therefore important to converge the Krylov subspace calculation as tightly as possible.

These results are also evidence for continuing to add residuals from converged solution vectors to the subspace, as they can contribute significantly to the solutions; new vectors generated are kept in the order of the solutions that generated them: the odd numbered new basis vectors are generated from the first solution and significantly contribute to the second solution, as shown.

1.4.4 Choosing number of starting vectors

For a non-orthonormal Krylov subspace basis, the rate of decay in the condition of the basis varies with the number of starting vectors, the choice of starting vectors, and the choice of preconditioner. The final value of the log base 10 of the inverse condition number of the scaled Gram matrix gives some sense of the quality of the final Krylov subspace basis, and is reported in 1.3. The more positive the log base 10 of the inverse condition number, the more orthogonal (and more stable) the basis.

The blank entries in table 1.3 are where the solver does not converge. Restarting after a fixed number of iterations may help – the solver does converge with an orthonormal Krylov subspace projection basis with all the options listed, without the need for restarts, with the root flip occurring in a similar fashion. (The inverse condition number is always 1 for an orthonormal Krylov subspace projection basis). Note that restarting too early would interfere with the “root-flip” required for convergence. Attempting the solve with null preconditioning or no preconditioning does not converge in 50 of iterations - with the “root-flip” not occurring

	number of starting vectors				
$\mathbf{A}^{[0]}$ for starting vectors		*no*	*no*	*yes*	*yes*
$\mathbf{A}^{[0]}$ for preconditioning		*yes*	*no*	*yes*	*no*
	2	-2.63510	-2.56132	-1.93622	-7.66052
	4	-2.27165	-2.83098	-3.33129	
	6	-2.55125	-2.88195	-2.14518	
	8	-1.63052	-1.63247	-1.68637	
	10	-1.62900	-1.50458	-1.45262	
	12	-1.64108	-3.30765	-1.42232	-5.44331

Table 1.3: Log 10 of the final inverse condition number of a non-orthonormal Krylov subspace approach with different numbers of starting vectors.

Using the Davidson preconditioning for solving for 2 lowest excitations of Uracil with PBE0. Different options are used for $\mathbf{M}$ for preconditioning and to determine start vectors.

in those 50 steps.

While determining the starting vectors with either $\mathbf{D}$ or $\mathbf{A}^{[0]}$ has no clear trend, it seems quite obvious that Davidson preconditioning with $\mathbf{A}^{[0]}$ is superior to Davidson preconditioning with $\mathbf{D}$ for this system with a non-orthogonal Krylov subspace basis, with the Krylov projection basis having a better condition. These results may change with different approximations taken in the matrix vector product, particularly if time is saved by directly multiplying $\mathbf{D}$ or $\mathbf{A}^{[0]}$ with the new vectors without scaling as part of the algorithm for obtaining the new matrix-vector products.

Avoiding “root-flips”

Converging more solutions does reduce the problem of “root-flips”. Figure 1.6 show the flexibility of *libkrylov* by varying the number of solutions calculated while keeping the number of starting vectors constant. The dotted lines show the convergence behavior with TURBOMOLE’s recommended number of starting vectors, which does show the same trend.

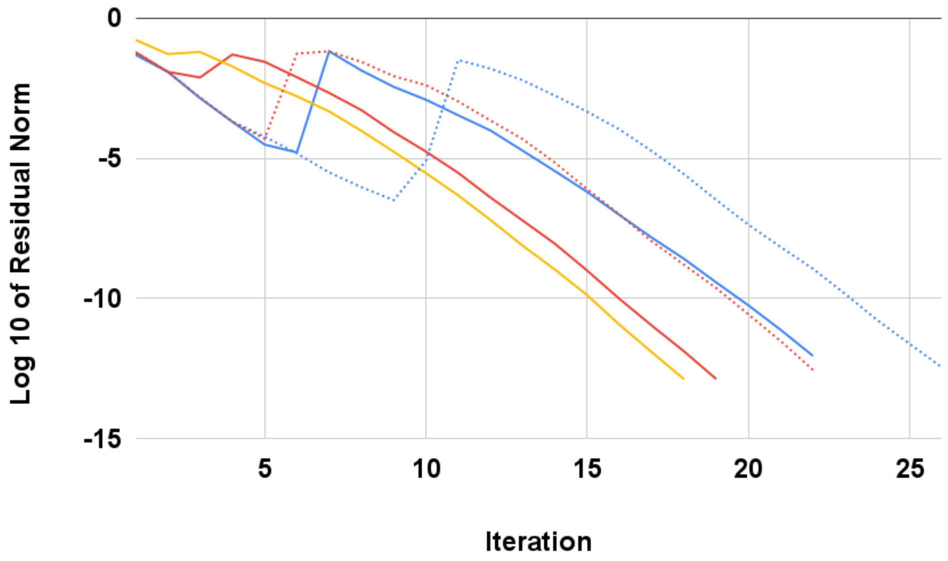

Figure 1.6: Krylov Subspace approach for solving Uracil with PBE0, using **D** for preconditioning and determining start vectors.
The solid lines are for 8 starting vectors, with blue for 2 solution, red for 3 solutions and yellow for 4 solutions. The dotted lines are for the same number of solutions, but with TURBOMOLE's recommended number of starting vectors which is twice the number of desired solutions.

1.4.5 Linear Dependence in preconditioning

With some options for Davidson preconditioning apparently leading to better convergence in general, and some options not leading to convergence, it is useful to try some of the more computationally expensive options in *libkrylov*.

Comparing $\mathbf{A}^{[0]}$ and D

Selecting the case where 8 starting vectors are chosen by using **D**, the Krylov subspace solve for the 2 lowest eigenvectors, with a non-orthonormal Krylov projection basis was attempted first with the Davidson preconditioner, then with all combinations of SVD of the new vectors and the first variant of Sleijpen preconditioning, for Uracil.

The residual norms behave very similarly with and without Sleijpen's and the SVD's reduction of linear dependence, shown in figure 1.7.

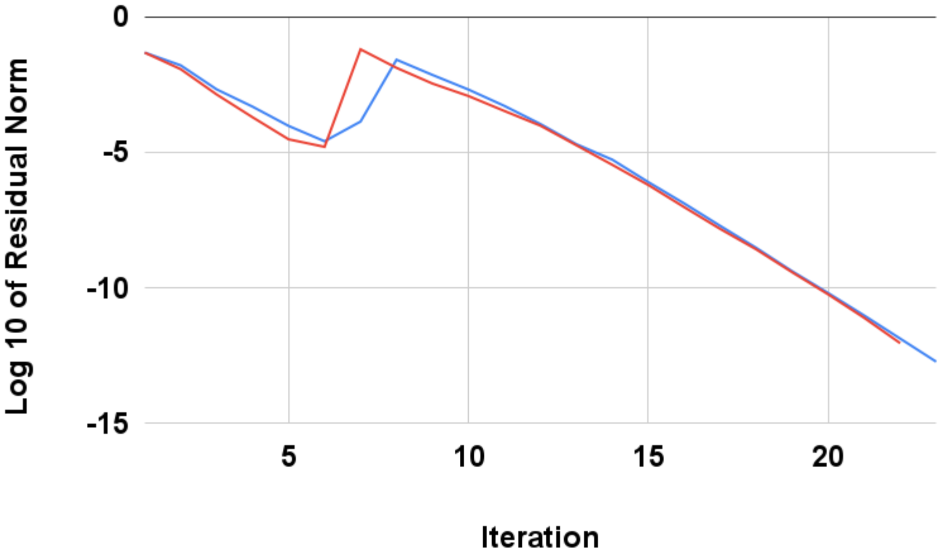

Figure 1.7: Krylov Subspace approach for solving Uracil with PBE0, using $\mathbf{D}$ for determining start vectors, with 8 starting vectors.
Using $\mathbf{A}^{[0]}$ for preconditioning is in blue, using $\mathbf{D}$ for preconditioning is in red. There is no significant difference between Davidson, Sleijpen's 1st variant, and additional SVD of new vectors once the above option is fixed.

The inverse condition number of the final scaled Gram matrix however shows differences, tabulated in table 1.4. These differences are possibly specific to this particular problem.

The condition of the Gram matrix is improved by SVD of new vectors, which can be important if the solve is facing numerical instability in the basis. Considering that the first variant of Sleijpen is useful to eliminate the projection of the residual onto the corresponding solution, a greater improvement in the condition with the use of the first variant of Sleijpen preconditioning indicates that the preconditioner matrix projected a greater magnitude of the residual onto the corresponding solution, implying that the preconditioner matrix is a better approximation of $\mathbf{A}$. Using $\mathbf{D}$ to construct the Davidson preconditioner appears to project residuals onto their corresponding solution more than using $\mathbf{A}^{[0]}$. This implies that for Uracil, for Davidson preconditioning, $\mathbf{A}^{[0]}$ seems to be the better choice for precondi-

Preconditioning	SVD of new vectors	Log 10 of final inverse condition number
Davidson with $\mathbf{A}^{[0]}$	No	-1.63052
Davidson with $\mathbf{D}$	No	-1.63247
Sleijpen 1st with $\mathbf{A}^{[0]}$	No	-1.63089
Sleijpen 1st with $\mathbf{D}$	No	-1.60974
Davidson with $\mathbf{A}^{[0]}$	Yes	-1.06854
Davidson with $\mathbf{D}$	Yes	-1.40528
Sleijpen 1st with $\mathbf{A}^{[0]}$	Yes	-1.07229
Sleijpen 1st with $\mathbf{D}$	Yes	-0.84456

Table 1.4: Using various preconditioning tools and and non-orthonormal Krylov subspace approach for solving for 2 lowest excitations of Uracil with PBE0 with 8 starting vectors determined from $\mathbf{D}$

tioning compared to $\mathbf{D}$ because it is a worse approximation to $\mathbf{A}$. *For other systems and parameters where* $\mathbf{M}$ *chosen may be a very good approximation for* $\mathbf{A}$ *in projecting residuals onto solution vectors, Sleijpen's Jacobi-Davidson preconditioning may be very useful.*

Where Davidson does not converge

From table 1.3, it can be seen that using $\mathbf{D}$ for Davidson preconditioning, and $\mathbf{A}^{[0]}$ for start vectors, for solving for the lowest two eigenvectors with a starting subspace of six, with a non-orthonormal Krylov subspace basis, the solve does not converge.

Thankfully, with a few other combinations of options, that are available in *libkrylov* the solver does converge, with the convergence behavior shown in figure 1.8 and figure 1.9.

If convergence is hindered by the preconditioner significantly projects the residual onto its corresponding solution, as it does in this case, eliminating that linear dependence with the first variant of Sleijpen preconditioning allows the solver to converge. An SVD of new vectors, where the linear dependence of Sleijpen preconditioned residuals among themselves is reduced, still results in a smaller condition number of the scaled Gram matrix. (Interestingly, just SVD of the new vectors does not allow the solver to converge, indicating that it is the linear dependence between the preconditioned residual and the corresponding solution that

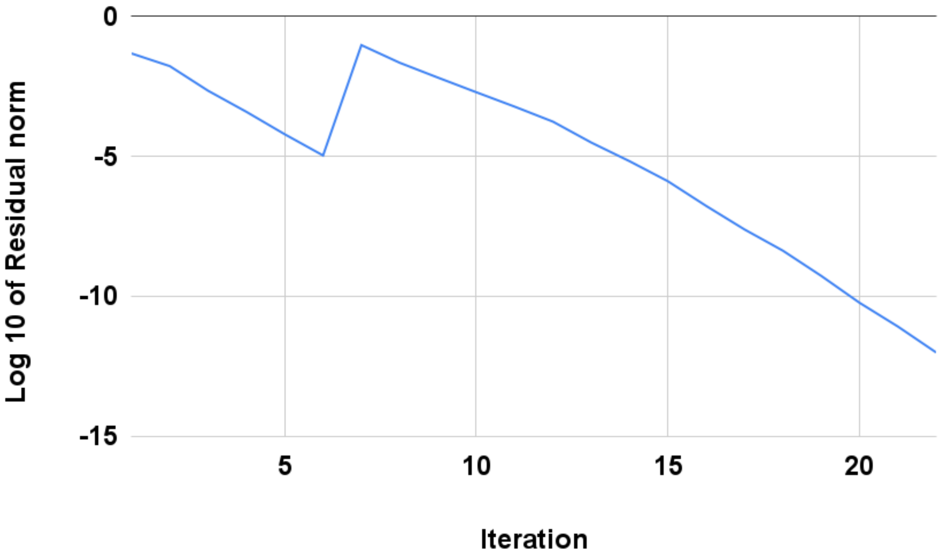

Figure 1.8: Krylov subspace solve: Decay of Residual Norm over iterations for calculating two lowest eigenvectors, with 6 starting vectors, determined with $\mathbf{A}^{[0]}$, preconditioning using $\mathbf{D}$
For the following methods: Davidson preconditioning with an orthonormal Krylov subspace projection, the first variant of Sleijpen preconditioning with an orthonormal Krylov subspace projection, first variant of Sleijpen preconditioning with a non-orthonormal Krylov subspace projection, and first variant of Sleijpen preconditioning with a non-orthonormal Krylov subspace projection with an SVD of new vectors.

is interfering with the algorithm.) It is useful to note that the residuals from iteration 7 and 8, where the residuals have greater norms due to "root-flipping", and are the residuals where the SVD of the new vectors has the greatest benefit.

1.5 Conclusions

Libkrylov presents more options in attempting a Krylov subspace solve, which allows for flexibility in the solve if features of the problem are identified, features that can be identified with functions that are part of the design of *libkrylov*.

For the non-orthonormal Krylov subspace projection basis, a better approximation of the

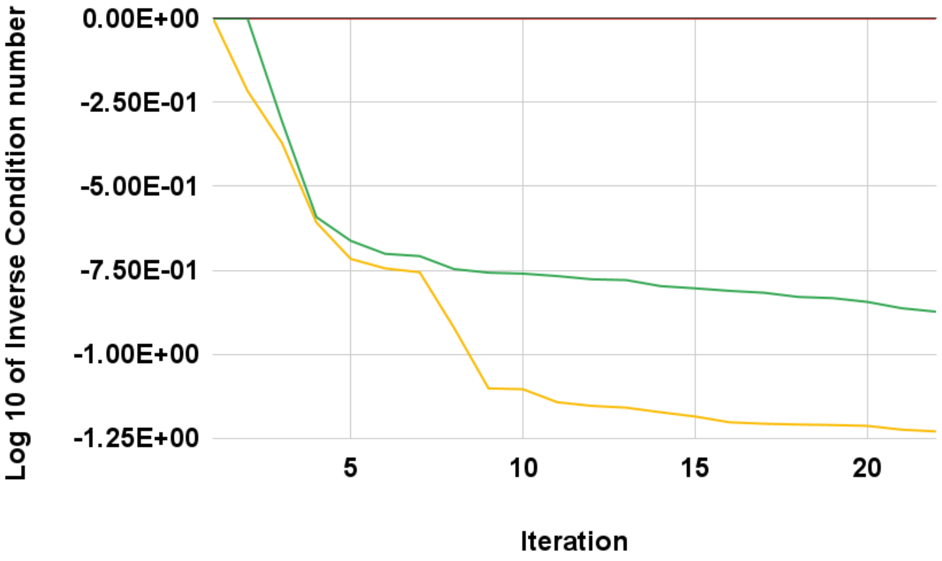

Figure 1.9: Krylov subspace solve: Decay of Inverse condition number over iterations with various methods following figure 1.8
For the following methods: first variant of Sleijpen preconditioning with a non-orthonormal Krylov subspace projection(yellow), and first variant of Sleijpen preconditioning with a non-orthonormal Krylov subspace projection with an SVD of new vectors(green).

coefficient matrix of the problem for preconditioning can project the residual back on the corresponding solution, leading to linear dependence of the Krylov subspace basis. In these cases, a specific variant of Sleijpen's Jacobi-Davidson preconditioning can be useful to reduce this linear dependence; and can be further improved by an SVD of new vectors.

Chapter 2

Spontaneous Emission rates from Tully Surface Hopping

The material in this chapter is based upon work supported by the National Science Foundation under OAC-1835909, and the Department of Energy under under DE-SC0018352.

All calculations were performed with a custom version of the TURBOMOLE 7.6 package (https://www.turbomole.org/).

2.1 Introduction

Since Tully's publication of his method,[26] Tully's Fewest Switches Surface Hopping (FSSH) has been applied to a wide range of photochemical processes involving electronically excited states. [13, 27, 28, 29, 30, 31] Each trajectory in Tully Surface Hopping is propagated independently and iteratively. A trajectory consists of finite time-step Born-Oppenheimer classical nuclear dynamics alongside a model quantum probability density propagation. The model quantum density is propagated with a model Hamiltonian containing the energy difference of the electronic states and the coupling of the electronic states through the classical nuclear velocity, which is $\mathbf{v} \cdot \mathbf{d}_{jk}$ for state j and state k. A swarm of trajectories are launched, experiencing only one electronic state's (the active state) energy and forces; and can change their active state according to a probability calculated from the model quantum density that is propagated alongside each classical trajectory. The resulting trajectories are interpreted *together* as an ensemble, using and requiring ergodicity within each trajectory and between trajectories.

A challenge in these calculations is obtaining multiple excited states and their properties for each step of the trajectory, with many previous applications focused on one excited state strongly coupled to another state of the chemical system. Turbomole's implementation of Tully surface hopping in an adiabatic basis of states has previously been extended to account for multiple excited states and *all* non-adiabatic coupling between states considered, and applied to systems where there is strong coupling between the second and first excited states of closed-shell systems.[27]

Applying Turbomole's implementation of Tully Surface Hopping to predict the emission spectra of molecules without assuming the emissive state is a logical follow-up application. Ideally, a "first-principles" approach where the electromagnetic field is treated explicitly is favored. However, the oscillator strengths f_k and excitation wavelengths λ_k are already calculated at each time step, allowing for a simple calculation of Einstein coefficients for spontaneous emission from the active state to the ground state, and an intensity count by multiplying by the time-step size, for active state k to ground state 0,

$$A_{k0}(t_i) = 2\pi \frac{e^2 f_k(t_i)}{\epsilon m_e c(\lambda_k(t_i))^2} \tag{2.1}$$

$$I_{k0}(t_i) = A_{k0}(t_i)\Delta t N_A \tag{2.2}$$

which for a single trajectory can be summed over a time interval where the trajectory has only one active state.[32]

The isomers of Azulene provide a good test set for this application, despite being weakly emitting, as there is a consistent molecular formula among them (thus consistent nuclear masses and Cartesian degrees of freedom) and clear differences in their emission spectra. Azulene is the textbook example of an anti-Kasha molecule,[33, 34] with no emission from the first singlet excited state observed without exciting directly to the first singlet excited state. [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49] The other isomers are expected to have predicted emission spectra dominated by emissions from the first singlet excited state, in line with Kasha's rule.[50]

However, initial attempts with this application were hindered odd behavior of trajectories, which can be traced back to (1) a combination of different chemical features of the chemical system under study from previous applications, (2) related features of the Tully Surface Hopping method and (3) subsequent choices in the implementation of Tully Surface Hopping. These problems are not well considered in the literature.

This chapter begins with describing the computational parameters and the chemical system under study, and distinguishing these fused ring system from the chemical systems of previous investigations with TURBOMOLE. This provides a base-line for analyzing the Tully Surface Hopping method analytically, describing where the method might not be suitable, to guide later analysis of the trajectories. While the Tully surface hopping algorithm is relatively straightforward (described explicitly in full in the appendix B.1), additional approximations were used in the previous implementation, and they are briefly discussed.

The trajectories are then presented and analyzed, which, despite the errors in the propagation, appear to predict Kasha's Rule qualitatively.

2.2 Computational Details

Azulene, Bicyclo[6.2.0]decapentaene, and Napthalene are examined using the electronic structure code and molecular dynamics code as implemented in a custom version of TURBOMOLE 7.6.

Kohn Sham Density Functional Theory is used to determine the ground state electronic structure and properties. The hybrid GGA functional PBE0 is used,[24, 25] with gridsize 4 for evaluating the gradient of the density numerically, and def2-SVP basis set[23] The scf energy convergence threshold was set to 10^{-8} Hartrees, and one electron reduced density matrix convergence threshold was set to 10^{-8} as well. The resolution of identity (RI-J) for the J term is used.[51]

The ground state gradient with respect to nuclear displacement is evaluated analytically.

Ground state *ab-initio* molecular dyanamics was launched for all molecules, from the ground state minimum energy geometry with a random nuclear velocity corresponding to either 300K or 500K of temperature. With a time step of 20 a.u. (which is approximately 0.5 fs), 500 steps were used for the system to thermally equilibrate, then the geometry and velocity is sampled every 100 steps to initiate non-adiabatic molecular dynamics. The classical dynamics is done with the leapfrog algorithm. [52] The energy drift of the ground state molecular dynamics is under 0.4 miliHartree over 20000 steps.

Excitation	Dominant Occupied	Dominant Virtual	ω (wavenumbers)	f
$S_0 \to S_1$	HOMO	LUMO	20339.7	0.012654
$S_0 \to S_2$	HOMO HOMO-1	LUMO+1 LUMO	31236.2	0.002029
$S_0 \to S_3$	HOMO-1 HOMO-2	LUMO+1 LUMO	39645.3	0.062073
$S_0 \to S_4$	HOMO-3	LUMO	47542.8	0.000138

Table 2.1: Excitations at geometry minimum for Azulene
, with the parameters described in the main text

The Tamm Dancoff Approximation and adiabatic approximation are used for the linear response calculation and excited state gradient correction.[14] All excited state gradients and all non-adiabatic couplings are calculated for the 4 lowest excitations considered. The pseudo-wavefunction approximation is used for the non-adiabatic coupling.[53]

5 trajectories for each molecule and temperature were launched. With the same time step of 20 a.u., the cycle number of the trajectory can be multiplied by the time step to give a time index for quantities below.

For these systems, spectra and Kasha/non-Kasha behavior is usually observed in solvent, and solvation is known to change excited state properties. However, besides limitations of the implementation, the micro-canonical ensemble of Tully surface hopping is not immediately compatible with a solvent heat bath. The dynamics is run without solvent correction.

2.2.1 Features of molecular system

Carrying out single point excitations from the ground state geometry of these molecules, their excited state character (at that geometry) can be characterized, in tables 2.1,2.2 and 2.3.

Excitation	Dominant Occupied	Dominant Virtual	ω (wavenumbers)	f
$S_0 \rightarrow S_1$	HOMO	LUMO	22272.4	0.014712
$S_0 \rightarrow S_2$	HOMO HOMO-1	LUMO+1 LUMO	25441.3	0.000002
$S_0 \rightarrow S_3$	HOMO-1 HOMO-2	LUMO+1 LUMO	39670.8	0.126145
$S_0 \rightarrow S_4$	HOMO-3	LUMO	46819.1	0.000252

Table 2.2: Excitations at geometry minimum for Bicyclo[6.2.0]decapentaene , with the parameters described in the main text

Excitation	Dominant Occupied	Dominant Virtual	ω (wavenumbers)	f
$S_0 \rightarrow S_1$	HOMO HOMO-1	LUMO+1 LUMO	37160.3	0.000000
$S_0 \rightarrow S_2$	HOMO HOMO-1	LUMO LUMO+1	38225.6	0.081623
$S_0 \rightarrow S_3$	HOMO HOMO-2	LUMO+2 LUMO	47884.1	0.000000
$S_0 \rightarrow S_4$	HOMO-2 HOMO-1	LUMO+1 LUMO+2	52149.6	0.000000

Table 2.3: Excitations at geometry minimum for Naphthalene , with the parameters described in the main text

2.3 Examining Tully Surface Hopping

Tully surface hopping, as close as possible to Tully's recommendation,[26] is used for this investigation.

2.3.1 Notation and quantities

The time step size is denoted with Δt.

For most iterations, indexed by a time t_i, there is data from the previous iteration: the active state k, the classical nuclei position $\mathbf{x}(t_i)$, classical nuclei velocity $\mathbf{v}(t_i - \frac{\Delta t}{2}) = \mathbf{v}(t_{i-1} + \frac{\Delta t}{2})$, and model quantum nuclei probability density $\boldsymbol{\rho}(t_i - \frac{\Delta t}{2})$. The diagonals of $\boldsymbol{\rho}$ are referred to as the populations, and the off-diagonals as the coherences. The decision to hop is also determined on the previous step – the Tully probability at t_i is interpreted as the probability of switching in the interval $[t_i, t_{i+1})$, thus each time step makes the decision if a hop is to be attempted on the next step. The velocities are at half time-step offsets due to the leapfrog verlet algorithm, which is a symplectic propagator and expected to preserve time reversal symmetry. The model quantum density is also at half time-step offset for time reversal symmetry.

An adiabatic basis is chosen for the electronic states, with the number of states considered being a parameter in the calculation. The classical nuclei positions $\mathbf{x}(t_i)$ are used to calculate the Born-Oppenheimer(BO) potential energies $\mathbf{V}(t_i)$, forces $\mathbf{F}(t_i)$, and non-adiabatic coupling $\mathbf{d}(t_i)$ for that time step; for all possible electronic states and between all electronic states. The energy of the ground state is set at zero, which allows for the excitation energy from the ground state to the excited state p, ϵ_p, to be used for the potential energy of each state. $V_{pq} = \epsilon_p \delta_{pq}$. $\mathbf{d}_{pp} = \mathbf{0}$ due to the adiabatic basis. The forces on all states should be calculated, as

$$-\left.\frac{\partial V_{pp}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}(t_i)} = \mathbf{F}_p(\mathbf{x}(t_i))$$

The averaged velocity $\mathbf{v}(t_i)$ can be calculated by averaging $\mathbf{v}(t_i + \frac{\Delta t}{2})$ and $\mathbf{v}(t_i - \frac{\Delta t}{2})$.

Model Hamiltonian

The model quantum Hamiltonian $\mathbf{H}(t_i)$ is thus constructed as

$$\mathbf{H}(t_i) = \mathbf{V}(t_i) - i\mathbf{v}(t_i) \cdot \mathbf{d}(t_i)$$

and is used to construct a propagator

$$\mathbf{U}(t_i) = \exp(-i\mathbf{H}(t_i)\frac{\Delta t}{2})$$

When classical velocities are changed instantaneously at a time step (during a hop), there are multiple velocities associated with the same time (t_i). Each set of classical velocities can be associated with and indexed by an adiabatic electronic state k.

Since the Hamiltonian also uses these classical velocities, there are distinct Hamiltonians for each set of classical velocities at time (t_i). These Hamiltonians can likewise be indexed by adiabatic electronic states:

$$\mathbf{H}_k(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_k(t_i) \cdot \mathbf{d}(t_i)$$

This extra index is unfortunate, but is required to properly track the time propagation.

Note there are two different kinds of indexes - one index over the electronic states, running from the ground or reference state zero to the highest excited state considered, and an index over the 3 Cartesian degrees of freedom for each classical nuclei. Some objects like $\mathbf{d}$ are hence matrices of vectors, where $\mathbf{d}_{pq}$ is a vector over the Cartesian degrees of freedom all nuclei, for the non-adiabatic coupling vector between state p and q. The mass $\mathbf{m}$ is indexed over the Cartesian degrees of freedom, but the degrees of freedom attached to the same nuclei have the same mass.

Hopping probabilities

Using the model quantum density as an approximation to the nuclear-electron quantum density (as Tully does when comparing Surface Hopping to quantum calculations), and looking at $-i\mathbf{H}$ as the instantaneous rate of change of the model quantum density, the instantaneous rate of change in population from state k to state j is b_{jk}, defined as:

$$b_{jk}(t_i) = 2\,\mathrm{Real}(\rho_{jk}(t_i)\mathbf{v}_k(t_i) \cdot \mathbf{d}_{jk}(t_i)) \tag{2.3}$$

and the first-order change in population for time interval t_i, g_{kj} is defined as:

$$g_{kj} = \frac{b_{jk}(t_i)}{\rho_{kk}(t_i)}\Delta t \tag{2.4}$$

This first-order change in population is interpreted as a hopping probability from state k to state j, and set to zero if negative.

As Tully suggested,[26] and expanded by Parker and coworkers, [27] the goal is not to hop only to the state with the highest probability. Hence the probabilities of hopping to different states are summed cumulatively - assuming that the probabilities are normalized, using the cumulative sum of probabilities to hop will result in a distribution of classical active states matching the quantum populations.

One advantage of defining hopping probabilities this way is that the hopping probabilities are independent of the type of quantum propagator used.

Hopping events

Since the hopping probability is determined for an instantaneous time value instead of a time interval, there is a free choice of the time interval that the probability corresponds to, as long as it contains that time value and the hop occurs at the end of that time interval.

In switching from state k to state j, as the trajectories are part of the microcanonical ensemble with conserved total energy, the energy change in the electronic states must be matched by an opposing change in the classical nuclei kinetic energy.

As described by Tully,[26] published by Barbatti, [54] the classical velocities can be rescaled to carry out this conservation of total energy:

$$\frac{1}{2\mathbf{m}}\mathbf{d}_{jk}(t_i) \cdot \mathbf{d}_{jk}(t_i)\gamma^2 + \mathbf{v}_k(t_i) \cdot \mathbf{d}_{jk}(t_i)\gamma + \epsilon_j(t_i) - \epsilon_k(t_i) = 0 \tag{2.5}$$

$$\mathbf{v}_j(t_i) = \mathbf{v}_k(t_i) + \gamma_{\text{min}} \frac{\mathbf{d}_{jk}(t_i)}{\mathbf{m}} \tag{2.6}$$

There are two roots for this equation, and following convention the smaller magnitude is chosen. If there are no real roots γ, the hop is forbidden.

The full Tully surface hopping algorithm is reported in appendix B.1.

2.3.2 Time reversal symmetry

Given the interpretation of Tully Surface Hopping trajectories as a micro-canonical ensemble of constant total energy, the classical and quantum propagators should have time reversal symmetry, i.e., be symplectic propagators.[52] The hopping probabilities should also satisfy time reversal symmetry. A feature of this property is that the error in total energy for the classical trajectory should oscillate around zero, with minimal energy drift over the course of the trajectory. Conserving total energy is directly related to time translational symmetry by Noether's Theorem, which is important for consistent comparability of different time regions of each trajectory. The propagators used and calculation of hopping probabilities for this application are examined. This examination is specific to this particular implementation.

Model quantum propagator

The quantum propagator used is the first term of the Magnus expansion using the model Hamiltonian. A propagator $\mathbf{U}$ should have two arguments for its time interval, but for simplicity we keep only one argument of the time step:

$$\mathbf{U}(t_i) = \mathbf{U}\left(t_i, t_i - \frac{\Delta t}{2}\right) = \mathbf{U}\left(t_i + \frac{\Delta t}{2}, t_i\right)$$

While the Magnus expansion allows for an approximation of $\mathbf{U}(t)$ using a power series:

$$\begin{aligned}\mathbf{U}(t_2,t_1) &= \exp(-i\int_{t_1}^{t_2}\mathbf{H}(t)dt) \\ &= \exp(\Omega_1+\Omega_2+\Omega_3+...) \qquad (2.7)\\ \Omega_1(t_2,t_1) &= -i\int_{t_1}^{t_2}dt\mathbf{H}(t) \qquad (2.8)\\ \Omega_2(t_2,t_1) &= -\frac{1}{2}\int_{t_1}^{t_2}dt\int_{t_1}^{t}dt'[\mathbf{H}(t),\mathbf{H}(t')] \qquad (2.9)\\ \Omega_3(t_2,t_1) &= \frac{i}{6}\int_{t_1}^{t_2}dt\int_{t_1}^{t}dt'\int_{t_1}^{t'}dt''[\mathbf{H}(t),[\mathbf{H}(t'),\mathbf{H}(t'')]]+[\mathbf{H}(t''),[\mathbf{H}(t'),\mathbf{H}(t)]] \qquad (2.10)\end{aligned}$$

using square brackets for the commutator. It has been noted by Blanes [55] that the integral in $\Omega_1(t_2,t_1)$ still needs to be approximated and introduces an error convoluted with the error of the Magnus expansion. There are other ways to do a Magnus expansion to eliminate or extract a $\mathbf{H}_0$ from the expansion; but in this Tully model system, there is no suitable time independent component, since the energies of the states is also time dependent. What is known is $\mathbf{H}(t_i)$. Writing in our notation, for the half time step offset:

$$\begin{aligned}\Omega_1\left(t_i+\frac{\Delta t}{2},t_i-\frac{\Delta t}{2}\right) &= -i(\Delta t)\mathbf{H}(t_i)-i(\Delta t)^3\frac{1}{24}\left.\frac{\partial^2\mathbf{H}(t)}{\partial t^2}\right|_{t=t_i}+\mathcal{O}((\Delta t)^5) \qquad (2.11)\\ \Omega_2\left(t_i+\frac{\Delta t}{2},t_i-\frac{\Delta t}{2}\right) &= i(\Delta t)^3\frac{1}{12}\left[\mathbf{H}(t_i),\left.\frac{\partial\mathbf{H}(t)}{\partial t}\right|_{t=t_i}\right]+\mathcal{O}((\Delta t)^5) \qquad (2.12)\end{aligned}$$

which, if we truncate Ω to second order in Δt to match the classical propagation, leaves the quantum propagator described above, without the need to numerically approximate the differentials of the Hamiltonian. While there are alternative Magnus expansions that do not exponentiate Ω to avoid these errors, these alternative Magnus expansions are not the options suggested by Blanes based on his examination of the formalism.[55]

The quantum propagation still needs to be done in half steps i.e.

$$\mathbf{U}\left(t_i + \frac{\Delta t}{2}, t_i - \frac{\Delta t}{2}\right) = \mathbf{U}\left(t_i + \frac{\Delta t}{2}, t_i\right)\mathbf{U}\left(t_i, t_i - \frac{\Delta t}{2}\right)$$

since the density at the full time step is required for calculating hopping probabilities.

With different Hamiltonians at the full time step during a hop at time t_i, the corresponding propagators should be considered to propagate from $t_i - \frac{\Delta t}{2}$ to t_i and from t_i to $t_i + \frac{\Delta t}{2}$, which is described in Appendix B.1.3. This sudden switching of Hamiltonians may not be significant, since what is changing is the classical velocity entering the Hamiltonian, which is rescaled as described in equation 2.5 and the following equations.

Classical propagator

The leapfrog verlet algorithm is known to be symplectic. However, in the event of a hop, there is a sudden switch in the potential energy surface, which is a sudden switch in the forces on the classical nuclei and a rescaling of the classical nuclei velocity to conserve total energy. Both result in the forces being discontinuous and not being conservative; the rescaling of velocity corresponds to an instantaneous force at that time step.

The discontinuity in forces has been noted in previous publications as an important factor leading to chemical reactions in the system.[29, 31, 30] However, if this discontinuity is significant, the symplectic property would be lost.

Hopping probabilities

The definition of b_{jk} is such that $b_{jk} = -b_{kj}$ (making the $\mathbf{b}$ matrix antihermitian like the propagator $\mathbf{U}$).

As Tully noted,[26] the choice of using b_{jk} to define g_{kj} allows for a clear linear dependence of g_{kj} on Δt, allowing for variable time step size. However, to ensure that all time steps are comparable, especially across different trajectories, the same time step size should be used. This property is hence not that useful.

The coherence of the density matrix only appears in the expression of the instantaneous rate, but not in the propagator $\mathbf{U}$ used for propagating the density. In implementation, the result is that both the propagator and hopping probability are constructed separately in the same `do` loops.

Tully requires setting negative g_{kj} to zero.[26] g_{kj} being negative has a physical intepretation, corresponding to an increase in population of state k due to state j, leading to "negative probability of hopping", when the population of state k is increasing, transferred from the population of state j. This is explained as part of Tully's Fewest switches criteria, where trajectories are to stay on active states for a significant amount of time, instead of rapidly switching between active states and experiencing an Ehrenfest force from the active states. This design of Tully surface hopping, with the hopping probability compared against a random number, should result in the ensemble's averaged classical population transfer to have time reversal symmetry if the hopping probability satisfies some normalization. Normalization requires that all the probabilities of hopping must sum up to one.

A significant problem is that such normalization of probabilities is not possible, since there is no equivalent definition g_{kk} - the probability of staying on the same state. This is despite some normalization of the rates b_{jk}, which themselves are insufficient to preventing individual g_{kj} from exceeding one, especially when ρ_{kk} becomes small and Δt is relatively large. Once the sum of probabilities exceeds one, the cumulative sum of probabilities no longer distributes the classical populations matching the quantum populations.

For the algorithm to run at the cumulative probability step, g_{kk} is set to zero such that the loop effectively skips state k. An unfortunate result is that the Tully system attempts no hops only when the sum of all probabilities is less than the random number ξ, which could explain the large number of forbidden hops observed. Again, the cumulative sum of probabilities no longer distributes the classical populations matching the quantum populations in this case.

With the criteria for rescaling velocity during a hop, it is possible that there are no real roots to equation 2.5, which corresponds to insufficient nuclear kinetic energy in the direction of the coupling for the hop. This problem of forbidden hops is also mentioned by Tully. This breaks both time-reversal symmetry and the correspondence of the quantum populations to the classical populations. This is probably neglected in previous applications with TURBOMOLE where the trajectories that remain on the excited state are not as important. For predicting emission spectra, the correct excited state is important and the forbidden hops are examined for the trajectories.

An additional complication is forced hops. Due to numerical instability in the response calculation for excited state properties when the first excited state is near degenerate with the reference state (or near a degeneracy in the ground state), the calculation can crash or produce nonsensical forces. To avoid a crash, if the active state is 1, a hop to the first state is forced when instabilities are expected, comparing a threshold against ϵ_1. This can be justified as a purely numerical instability, which would resolve to quantum population transfer and a normal hop on the subsequent time step with higher precision numerical (or analytical) methods. Therefore, force hops should preserve Tully's condition of the classical populations matching the quantum populations and time-reversal symmetry, but requires careful selection of the threshold at which to force hops. These are not as significant a concern for predicting the emission spectra as the ground state does not emit, except that the quantum probability density still needs to be propagated correctly to be averaged over the ensemble. Previous implementations in (and thus applications with) TURBOMOLE did not distinguish these forced hops from 'normal' hops.

Checks on propagation methods

Based on the above, there are a few parameters that can be used to check the propagators are implemented as intended.

Since there is no previous step, there is no hop on the first step. Since the coherences of the density matrix are zero for the first step, there is also zero probability of hopping on the second step. Checking for these zero probabilities is a quick check that the propagation is started correctly.

The total energy of the trajectories over the iterations can be plotted, and is expected to have simple oscillations with the symplectic propagation.

2.4 Changes from previous implementations

The above analysis applies to this particular implementation of Tully surface hopping, with the behavior of previous implementations being slightly different due to further approximations.

These additional approximations do not have any effect if the non-adiabatic coupling is numerically zero, which may occur in some applications.

Previous TURBOMOLE implementations had used the propagator and calculated the hopping probabilities differently –

$$\rho\left(t_{i+1}-\frac{\Delta t}{2}\right)=\mathbf{U}(t_i)\rho(t_i)\mathbf{U}^{\dagger}(t_i) \tag{2.13}$$

$$b_{jk}(t_i)=2\,\mathrm{Real}\left(\rho_{jk}\left(t_{i+1}-\frac{\Delta t}{2}\right)\mathbf{v}_{avg}(t_i)\cdot\mathbf{d}_{jk}(t_i)\right) \tag{2.14}$$

$$\rho(t_{i+1})=\mathbf{U}(t_i)\rho\left(t_{i+1}-\frac{\Delta t}{2}\right)\mathbf{U}^{\dagger}(t_i) \tag{2.15}$$

which may not lead to significant differences for small time step sizes and number of steps, if there is only a single relatively short and strong region of coupling that trajectories of interest pass through. The main consequence is the breaking of time-reversal symmetry, and perhaps an increase in the coherences of the quantum density matrix.

For forced hops, the reordering of populations was only carried out for two states, i.e.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10} \tag{2.16}$$

which was observed to lead to negative populations after a forced hop.

The previous implementation also reordered equation 2.5, and solves it twice – for a hop at time t_i the following equations are solved:

$$\frac{1}{2\mathbf{m}}\mathbf{d}_{lk}(t_{i-1}) \cdot \mathbf{d}_{lk}(t_{i-1})\gamma'^2 + \mathbf{v}_k(t_{i-1}) \cdot \mathbf{d}_{lk}(t_{i-1})\gamma' + \epsilon_l(t_{i-1}) - \epsilon_k(t_{i-1}) = 0 \tag{2.17}$$

$$\frac{1}{2\mathbf{m}}\mathbf{d}_{lk}(t_{i-1}) \cdot \mathbf{d}_{lk}(t_{i-1})\gamma''^2 + \mathbf{v}_k(t_{i-1}) \cdot \mathbf{d}_{lk}(t_{i-1})\gamma'' + \epsilon_l(t_i) - \epsilon_k(t_i) = 0 \tag{2.18}$$

Besides a slightly different γ, and ignoring some bugs in the calculation, the hop is forbidden if either of the two solves have no real roots. There are hence more opportunities for a hop to be forbidden since this equation is checked twice. These forbidden hops were also silent in previous implementations, with no indicator that hops were forbidden.

It is possible that there is some error cancellation between these additional approximations and changes to Tully's formulation.

Previous applications, sampling from a ground state leapfrog molecular dynamics trajectory, have taken the classical velocity at a half time-step earlier, that is at $(t_0 - \frac{\Delta t}{2})$ instead of at t_0. With sufficient sampling, the averaged error in velocity is zero, but may account for slightly slower dynamics seen in this application.

2.5 Total energy of trajectories

To validate trajectories, the total energy over time for each molecule at each temperature is plotted.

The initial energy of the trajectories differ due to different excitation energies from ground state positions and velocity of approximately the same total energy. As would be expected, the trajectories at a higher temperature have more variation in their initial excitation energy.

More interesting is the discontinuity in total energy seen after hopping, which is noticable when plotted at this scale. This discontinuity occurs the step after a hop is carried out. The largest discontinuities appear for the S_1 to S_0 hops of Bicyclo[6.2.0]decapentaene – the S_2 to S_1 hops are barely noticable in contrast. The largest discontinuity is at the forced hop of the 2nd trajectory of Bicyclo[6.2.0]decapentaene at 500K. While Napthalene's trajectories have discontinuities from hopping from S_2 to S_1 and back, they are not as drastic as Bicyclo[6.2.0]decapentaene.

Trajectories with additional energy discontinuities were rerun; all discontinuities are at the time step after a hop.

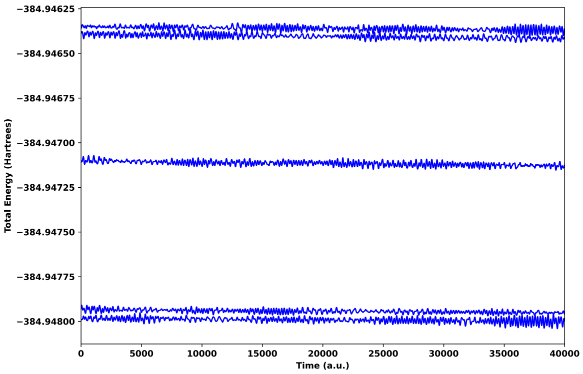

Figure 2.1: Total energies of Azulene trajectories at 300K

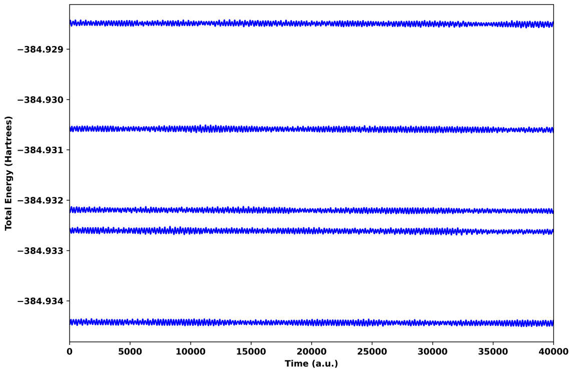

Figure 2.2: Total energies of Azulene trajectories at 500K

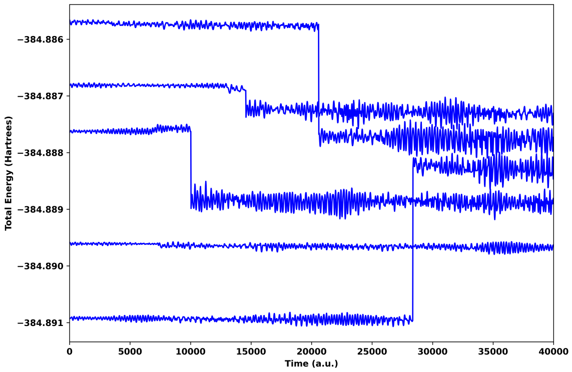

Figure 2.3: Total energies of Bicyclo[6.2.0]decapentaene trajectories at 300K

2.6 Hopping events

Traditionally, the average of classical trajectory's active state is used for analysis. Given only 5 trajectories for each molecule at each temperature, statistics on these trajectories may be misleading; thus the successful hopping events for each trajectory are listed with their probabilities, in table 2.4, 2.6 2.5, 2.7 and 2.8. The tables also highlights the probabilities being not normalized in Tully Surface hopping.

Based on the classical active state, Azulene molecules tend to remain on the S_2 state, Bicyclo[6.2.0]decapentaene hops relatively quickly to S_1 then tends to hop to S_0, while Naphthalene hops quickly to S_1, and hops for short periods of time back to S_2.

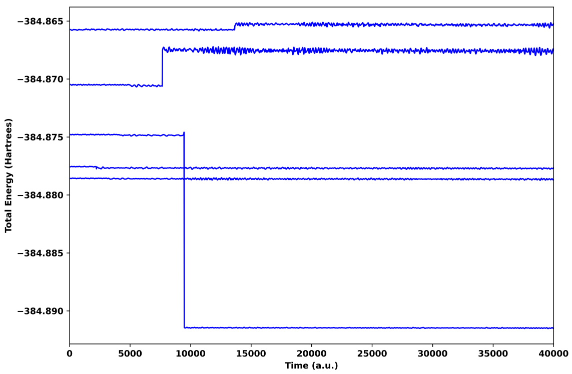

Figure 2.4: Total energies of Bicyclo[6.2.0]decapentaene trajectories at 500K

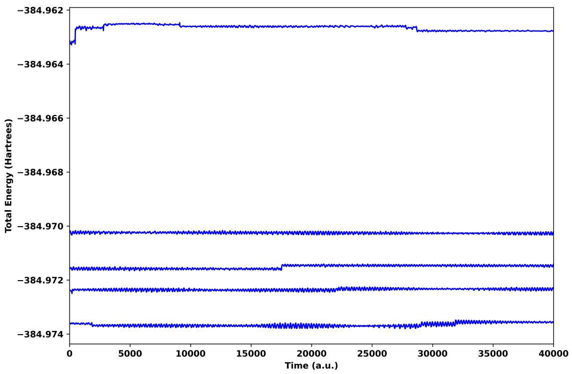

Figure 2.5: Total energies of Naphthalene trajectories at 300K

There is one Bicyclo[6.2.0]decapentaene trajectory at 500K, trajectory 2, which experiences a forced hop due to a small energy gap. This trajectory has the largest total energy discontinuity and is the trajectory with significant quantum S_3 and S_4 populations.

It is worthwhile to mention that these successful hops should be analyzed with the forbidden hops, which are quite frequent. A selected trajectory has the forbidden hops listed in appendix B.2.

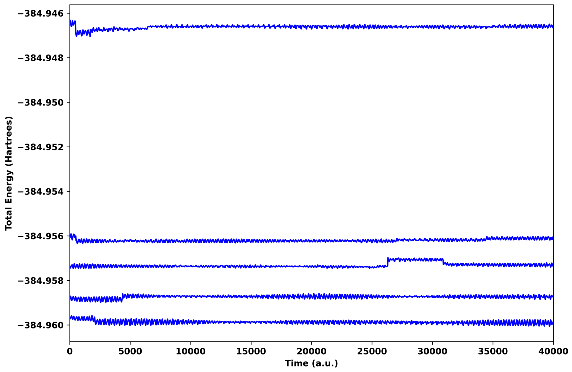

Figure 2.6: Total energies of Naphthalene trajectories at 500K

2.7 Examining Average populations

With different active states on different trajectories at the same time, in particular with hops up, a direct way to compare trajectories is to use the populations of the model quantum density.

Azulene's populations on S_2 remain close to one over all trajectories. Naphthalene's S_2 and S_1 populations oscillate for all trajectories. Some Bicyclo[6.2.0]decapentaene trajectories also oscillate between S_2 and S_1. Averaging across trajectories results in slower oscillations, reducing the magnitude of the S_3 and S_4 populations from the Bicyclo[6.2.0]decapentaene trajectory with the forced hop. Averaging also makes it difficult to distinguish the average populations from different temperatures – some details are averaged out. The averaged populations are shown in Figure 2.7,2.8, 2.9 and 2.10.

With a natural log plot not being very linear, an alternative is needed to attempt to qualitatively predict Kasha's rule.

Traj.	Event	Cycle	Source	Target	g_{kj}
1	hop	652	2	1	0.03277465638000
	hop	728	1	0	0.00335848831400
2	hop	365	2	1	0.02184184117000
3	hop	174	2	1	0.00322119114900
	hop	1029	1	0	0.00198280948300
4	hop	412	2	1	0.02090958591000
	hop	1418	1	0	0.00144891640400
5	hop	345	2	1	0.03325005826000
	hop	501	1	0	0.00021882655330

Table 2.4: Successful hops for all Bicyclo[6.2.0]decapentaene trajectories at 300K , labelled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

2.8 Predicted spectra

Assuming that the classical active states are a good indicator of the emitting state, the number of time steps corresponding to each active state is listed in table 2.9. These results qualitatively state Kasha's rule for Bicyclo[6.2.0]decapentaene and Naphthalene, but are quite different from the quantum populations. For Azulene, the results between the quantum populations and classical states agree that the molecule is anti-Kasha, because of poor coupling of the S_2 to S_1 state, where there is no hopping observed.

With the oscillator strengths of the active state at each time step, a histogram can be constructed with bins at every 100 wavenumbers. It is now possible to distinguish the electronic state that contributes to peaks in the emission spectra, without empirical weighting. For Bicyclo[6.2.0]decapentaene, it appears that the greater oscillator strength of the S_1 state contributes to that molecule's emission spectra being dominated by emissions from the S_1 state. For Naphthalene, the classical active state is also mostly on S_1 even though the quantum

Traj.	Event	Cycle	Source	Target	g_{kj}
1	hop	16	2	1	0.29257622810000
2	hop	14	2	1	1.70125097800000
	hop	1105	1	2	0.47595126020000
	hop	1106	2	1	0.35646770860000
3	hop	24	2	1	2.93343856000000
	hop	52	1	2	1.16353996400000
	hop	70	2	1	0.25733876470000
	hop	97	1	2	1.09904712100000
	hop	199	2	1	0.14121006040000
	hop	1257	1	2	0.71353021350000
	hop	1278	2	1	0.03333214509000
	hop	1386	1	2	0.80498932940000
	hop	1391	2	1	0.63921706300000
	hop	1415	1	2	3.36483697900000
	hop	1434	2	1	0.54110386420000
	hop	1461	1	2	0.34891792630000
	hop	1492	2	1	0.27774544290000
4	hop	16	2	1	0.27573446250000
5	hop	22	2	1	0.05049869268000

Table 2.5: Successful hops for all Naphthalene trajectories at 300K
, labeled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Traj.	Event	Cycle	Source	Target	g_{kj}
1	hop	249	2	1	0.32750995060000
	hop	383	1	0	0.00313626934800
2	hop	204	2	1	0.02331251347000
	Forced hop	473	1	0	−0.14362441970000
3	hop	112	2	1	13.28061337000000
4	hop	63	2	1	0.00260046371200
	hop	682	1	0	0.00824026965800
5	hop	162	2	1	0.01255217323000

Table 2.6: Successful hops for all Bicyclo[6.2.0]decapentaene trajectories at 500K , labeled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

populations are more evenly divided.

2.9 Conclusions

A consistent implementation of Tully surface hopping better respects time-reversal symmetry, therefore allowing for a “first-principles” prediction of emission spectra.

There are however various problems in the trajectories due to the algorithm. First, the Tully surface hopping probabilities are unbound, and is inconsistent with the quantum propagation. Second, there are large discontinuities in the total energy after certain hops.

Despite these difficulties, using the classical active state allows for a qualitative prediction of Kasha’s Rule and analysis of dominant electronic states. Bicyclo[6.2.0]decapentaene and Napthalene predominantly emit from the S_1 state, adhering to Kasha’s Rule, while Azulene emits only from the S_2 state as expected. The difficulties may prevent extending this study

Traj.	Event	Cycle	Source	Target	g_{kj}
1	hop	18	2	1	0.56481178810000
	hop	93	1	2	1.40611997800000
	hop	104	2	1	0.70006665800000
	hop	1430	1	2	0.18149310770000
	hop	1434	2	1	2.06179903800000
2	hop	24	2	1	6.12266367600000
3	hop	24	2	1	0.37970363000000
	hop	56	1	2	1.10591555900000
	hop	85	2	1	5.87241229500000
	hop	112	1	2	1.83236403500000
	hop	132	2	1	3.49750881300000
	hop	184	1	2	15.26882739000000
	hop	217	2	1	2.65724922200000
	hop	268	1	2	0.08519069350000
	hop	270	2	1	1.01492211500000
	hop	322	1	2	0.46885180430000
	hop	328	2	1	0.01670531895000
	hop	555	1	2	0.28123488970000
	hop	580	2	1	0.32707438260000

Table 2.7: Successful hops for Naphthalene trajectories at 500K, 1 , labeled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Traj.	Event	Cycle	Source	Target	g_{kj}
4	hop	15	2	1	0.43119210940000
	hop	1027	1	2	0.12716726170000
	hop	1047	2	1	0.05287423183000
	hop	1239	1	2	0.04820117049000
	hop	1241	2	1	3.23063878300000
	hop	1272	1	2	0.62395731720000
	hop	1277	2	1	0.05770330269000
	hop	1350	1	2	2.36398298800000
	hop	1361	2	1	0.18237939740000
	hop	1535	1	2	1.93675168900000
	hop	1565	2	1	1.22402530200000
	hop	1677	1	2	0.36565898480000
	hop	1685	2	1	0.44331170900000
5	hop	27	2	1	3.37916434900000
	hop	228	1	2	0.78455964230000
	hop	272	2	1	1.27530699100000
	hop	451	1	2	1.19215452800000
	hop	474	2	1	0.05008405450000
	hop	1348	1	2	0.62239811100000
	hop	1359	2	1	0.35757580480000
	hop	1710	1	2	2.96887888700000
	hop	1727	2	1	0.79552496370000

Table 2.8: Successful hops for Naphthalene trajectories at 500K, 2 , labeled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Molecule	Temperature	S_0	S_1	S_2	S_3	S_4
Azulene	300K			10000		
Azulene	500K			10000		
Bicyclo[6.2.0]decapentaene	300K	4320	3727	1953		
Bicyclo[6.2.0]decapentaene	500K	4459	4746	795		
Naphthalene	300K		9706	294		
Naphthalene	500K		9586	414		

Table 2.9: Number of steps each set of trajectories spends on each electronic state

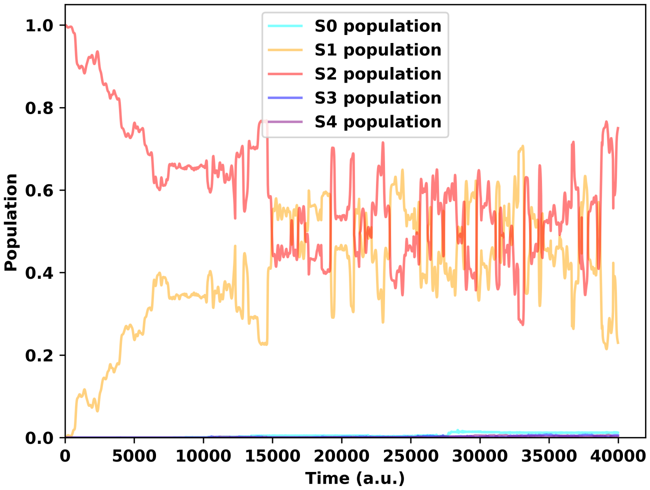

Figure 2.7: Average Quantum Populations of Bicyclo[6.2.0]decapentaene trajectories at 300K

to higher excited states and temperatures.

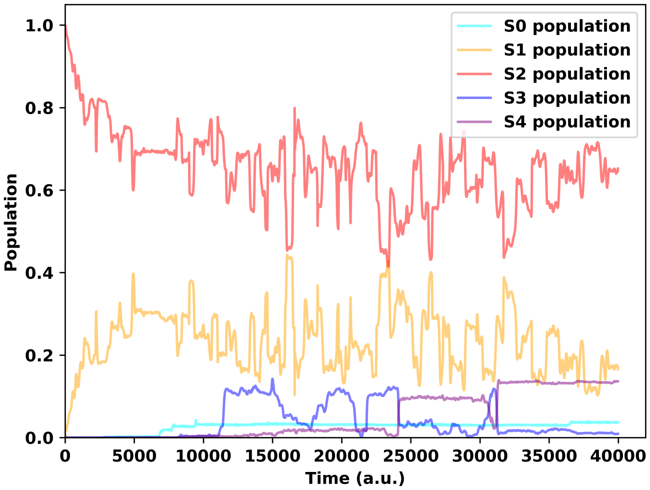

Figure 2.8: Average Quantum Populations of Bicyclo[6.2.0]decapentaene trajectories at 500K

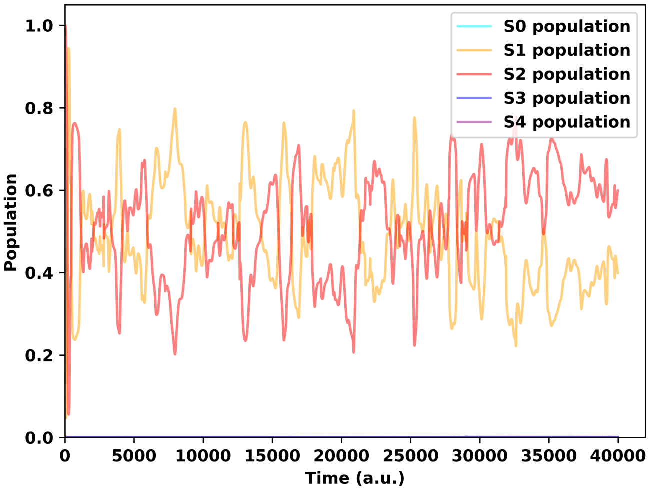

Figure 2.9: Average Quantum Populations of Naphthalene trajectories at 300K

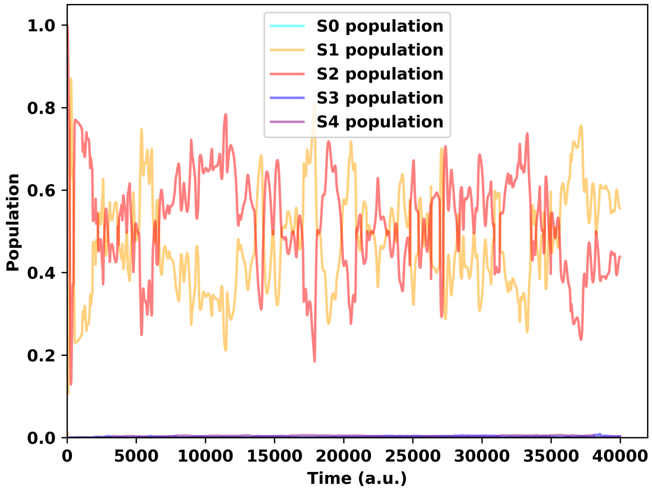

Figure 2.10: Average Quantum Populations of Naphthalene trajectories at 500K

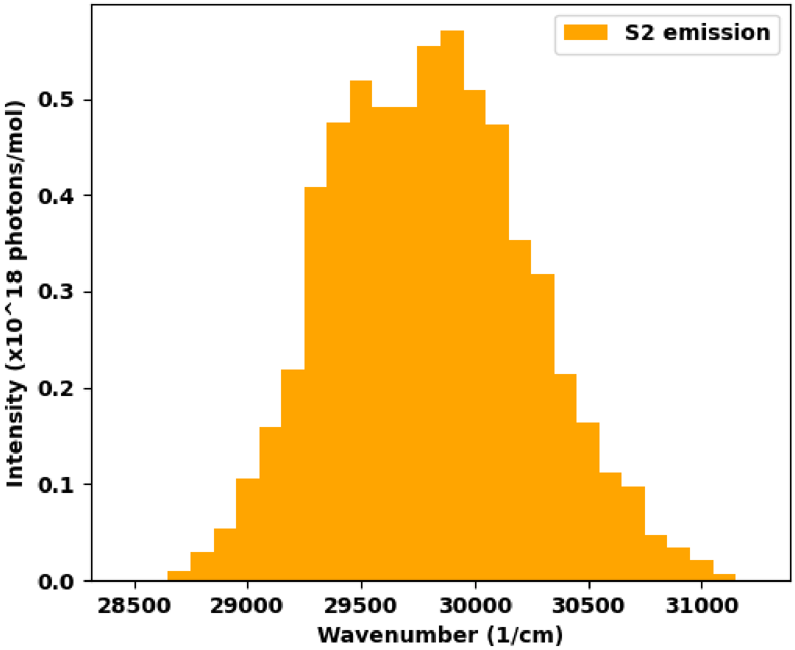

Figure 2.11: Predicted spectra for Azulene at 300K

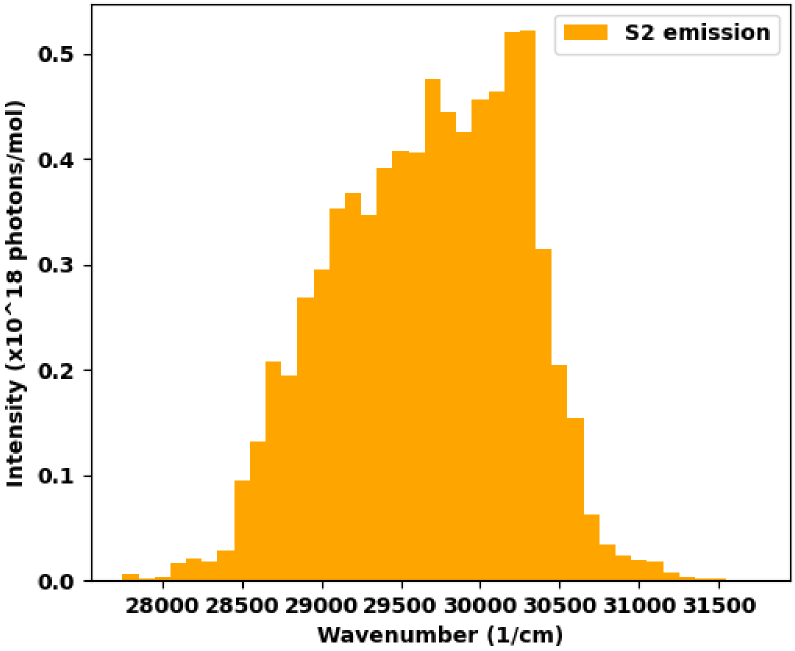

Figure 2.12: Predicted spectra for Azulene at 500K

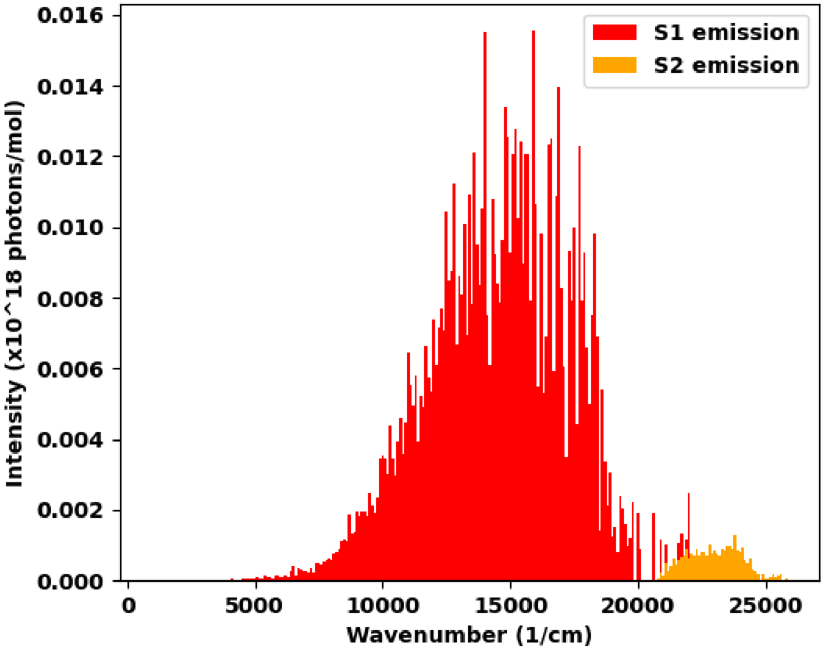

Figure 2.13: Predicted spectra for Bicyclo[6.2.0]decapentaene at 300K

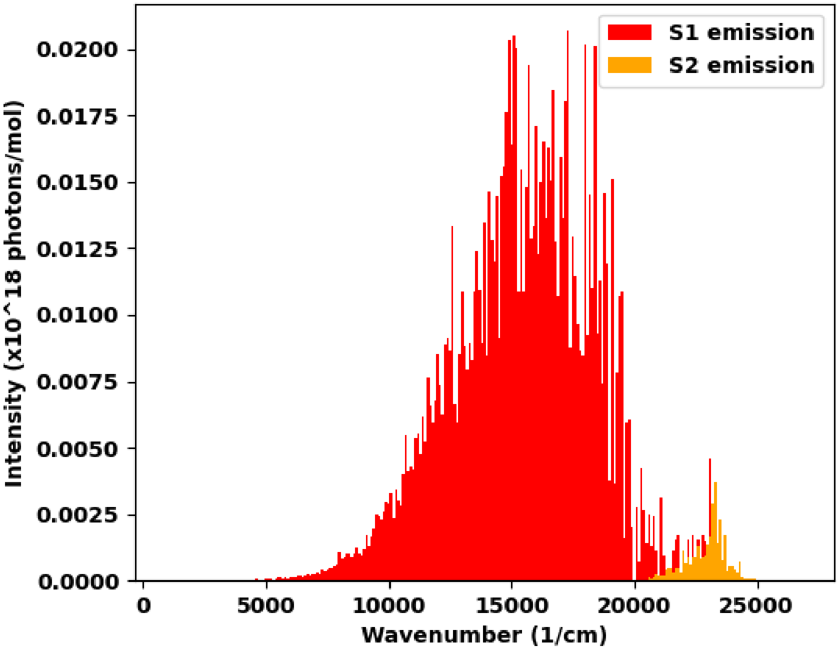

Figure 2.14: Predicted spectra for Bicyclo[6.2.0]decapentaene at 500K

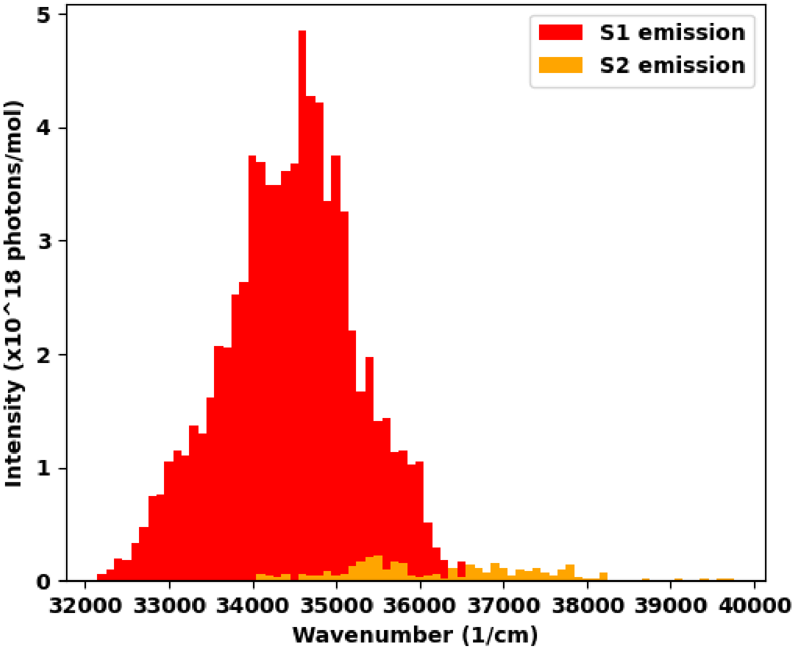

Figure 2.15: Predicted spectra for Naphthalene at 300K

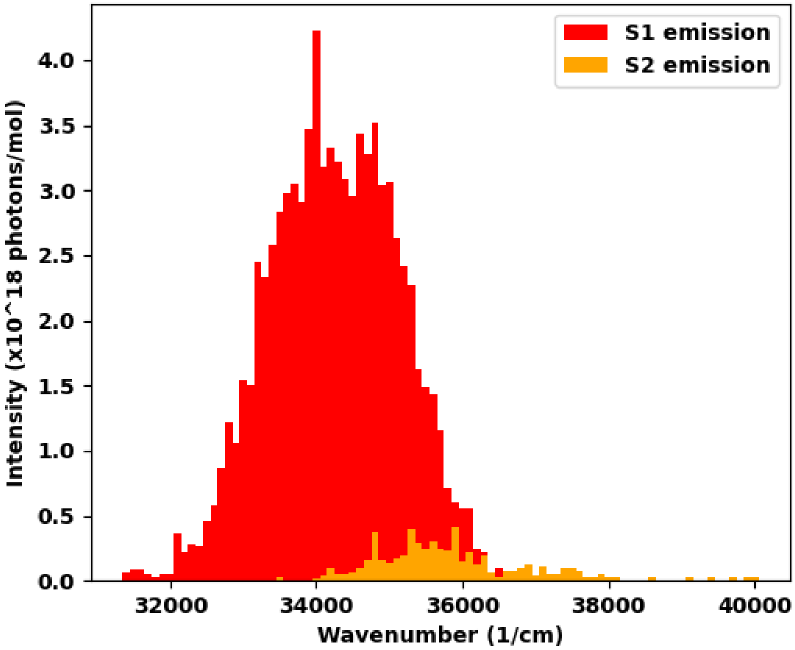

Figure 2.16: Time resolved spectra for Naphthalene at 500K

Chapter 3

Finite time-step Tully Surface Hopping

3.1 Future implementation

Changing the implementation in Turbomole to tackle the problems discussed in the previous chapter require first refactoring the code significantly before implementation of the new features below, as the previous implementation sought to separate the classical propagation and quantum propagation in sequential subroutines. With very large changes required for implementing these features, it is important to examine the features below analytically before an implementation is attempted. The first implementation should probably target a simpler code designed for model systems, ideally a model system with have multiple states and multiple degrees of freedom in the classical dynamics.

First, the Tully method is more closely examined to describe the errors of the model quantum system and the desired properties of the hopping probabilities.

A description of improved Tully Algorithms follows with an short analysis.

3.2 Analysis of Tully Model

3.2.1 Comparing the Tully Model system with a quantum nuclear-electron system

The instantaneous Tully surface hopping probability (equation 2.4) depends on both the Tully model Hamiltonian and the Tully model quantum density. It is reasonable to compare the Tully model Hamiltonian and quantum density to the nuclear-electron Hamiltonian and quantum density in a micro-canonical ensemble, as the surface hopping procedure is using the Tully model Hamiltonian and model quantum density to approximate the nuclear-electron Hamiltonian and nuclear-electron quantum density of the molecular system.[26] There is always a dependence of the electronic state on the nuclear states: in the Born-Oppenheimer separation, the electronic states and properties are implicitly dependent on the nuclear states; otherwise the electronic and nuclear states are not separable. Conversely, in the Born-Oppenheimer separation, the nuclear states are influenced by a potential derived from the electronic state.

The off-diagonals of the nuclear-electron Hamiltonian are the product of the coupling of the nuclear and electronic degrees of freedom in different states, which is complicated by the dependence of the electronic states on the nuclear states. The Tully model Hamiltonian approximates the off-diagonals of the nuclear-electron Hamiltonian by having one set of classical nuclear positions to determine the electronic states and one set of classical nuclear velocities which would otherwise be the coupling of nuclear states through a velocity operator. There are thus two practical approximations made in constructing the off-diagonals of the model Hamiltonian. The diagonals of the nuclear-electron Hamiltonian have a simpler relationship in that they should have the same value or energy(as the total energy is conserved); unlike the Tully model Hamiltonian where the states have energies that differ by the electronic excitation energies. This third "approximation" by Tully has an interesting consequence.

Combining all three approximations, the Tully model propagator defined from the Tully model Hamiltonian propagates the model quantum density with a phase factor θ_n which can be extracted by taking the natural log of $\mathbf{U}_{nn}(t_i)$:

$$\begin{aligned}
\rho_{jk}\left(t_i + \frac{\Delta t}{2}\right) &= \sum_{nm} \mathbf{U}_{jn}(t_i)\rho_{nm}(t_i)\mathbf{U}^{\dagger}_{mk}(t_i) \\
&= \rho_{jk}(t_i)\text{exp}\left(-i(\theta_j - \theta_k)\frac{\Delta t}{2}\right) \\
&+ \sum_{n \neq j} \mathbf{U}_{jn}(t_i)\rho_{nk}(t_i)\text{exp}\left(i\theta_k \frac{\Delta t}{2}\right) \\
&+ \sum_{m \neq k} \text{exp}\left(-i\theta_j \frac{\Delta t}{2}\right)\rho_{jm}(t_i)\mathbf{U}_{mk}(t_i) \\
&+ \sum_{n \neq j, m \neq k} \mathbf{U}_{jn}(t_i)\rho_{nm}(t_i)\mathbf{U}^{\dagger}_{mk}(t_i) \qquad (3.1)
\end{aligned}$$

where the phase factor θ_n cancels out some of the error from the practical approximations to the off-diagonals of the Hamiltonian. As the model quantum density is propagated iteratively, the incomplete error cancellation accumulates in the model quantum density; one of the contributors to the overcoherence of the model quantum density matrix. The instantaneous Tully Surface hopping probabilities are sensitive to this overcoherence, leading to the need for decoherence corrections in some systems.[56, 57]

Also as a result of this dependence on the model quantum density, the instantaneous Tully Surface hopping probabilities are implicitly dependent on the propagator used. While using the instantaneous Tully Surface hopping probabilities may make it simpler to switch between other approximate propagators, the implicit dependence of the instantaneous Tully Surface hopping probabilities on the propagator implies that using the propagator directly to determine hopping probabilites may be simpler.

3.2.2 Hopping Probabilities are between populations

The instantaneous hopping probabilities are the first order rate of change in the populations of the model quantum density matrix weighted by the population of the active state (equation 2.4). A replacement for the instantaneous hopping probabilities should also be determined by the change in populations or population-to-population transfer.

3.3 Normalized hopping probabilities

A remedy to the problems with the hopping probability, given that the propagator is explicitly calculated and the error analysis already presented for the propagator, is to use the propagator **U** directly to determine hopping probabilities. Note that the propagator is anti-hermitian.

The contribution to the population of state j at time t_i from the population of state k at t_{i-1} is:

$$b_{jk}(t_i) = U_{jk}(t_i, t_{i-1})\rho_{kk}(t_{i-1})U^{\dagger}_{kj}(t_i, t_{i-1}) \tag{3.2}$$

Notably,

$$\rho_{jj}(t_i) = b_{jj}(t_i) + \sum_{k\neq j} b_{jk}(t_i) + U_{jk}(t_i, t_{i-1})\rho_{kj}(t_{i-1})U^{\dagger}_{jj}(t_i, t_{i-1}) + U_{jj}(t_{i+1}, t_i)\rho_{jk}(t_i)U^{\dagger}_{kj}(t_{i+1}, t_i)$$

All the contributions of the population of state k at t_{i-1} to all other populations at t_i is

$$N_k(t_i) = \sum_j b_{jk}(t_i)$$

which is an ideal norm, since this b_{jk} is real and positive.

Normalized hopping away probability would be

$$g^{+}_{kj}(t_i) = \frac{b_{jk}(t_i)}{N_k(t_i)} = \frac{U_{jk}(t_i, t_{i-1})U^{\dagger}_{kj}(t_i, t_{i-1})}{\sum_s U_{sk}(t_i, t_{i-1})U^{\dagger}_{ks}(t_i, t_{i-1})} \tag{3.3}$$

In writing this definition of g_{kj}, it might appears necessary to include

$$U_{jk}(t_i, t_{i-1})\rho_{kj}(t_{i-1})U^{\dagger}_{jj}(t_i, t_{i-1}) + U_{jj}(t_i, t_{i-1})\rho_{jk}(t_{i-1})U^{\dagger}_{kj}(t_i, t_{i-1})$$

in the sum, which are related to Tully's original description(using coherences). That term is real, but can be negative. However, for the Tully classical dynamics, states evolve on one state at a time, and the correspondence is between the quantum populations and the classical populations. The coherences are not present in the classical dynamics, thus the hopping probability should only concern itself with population to population transfer. Tully's method is not designed to account for coherent states.

It is interesting that there is no dependence on the density matrix. This might solve part of the overcoherence or memory problem due to dependence on the density matrix (overcoherence due to the model Hamiltonian is NOT solved by this probability definition). More importantly, with multiple states, there is no increase is the probability of hopping to all others states when the population of the active state decreases.

The dependence of the probabilities on the time step is harder to see, buried in the definition of U. It is no longer straightforward to carry out error analysis with a variable time step. Howver, a consistent time step for all propagations must be used, which should be the case for the purposes of comparing all trajecotories equally.

However, Tully wanted a method with fewest switches, pointing out that it is undesired for two states coupled strongly to each other to hop rapidly back and forth.

The normalized hopping back probability, much like the de-excitation part of the excitation vector in Response Theory, is:

$$g^{-}_{kj}(t_i) = \frac{b_{kj}(t_i)}{N_j(t_i)} = \frac{U_{kj}(t_i, t_{i-1})U^{\dagger}_{jk}(t_i, t_{i-1})}{\sum_s U_{sj}(t_i, t_{i-1})U^{\dagger}_{js}(t_i, t_{i-1})} \tag{3.4}$$

Note that $g^{+}_{kk}(t_i) = g^{-}_{kk}(t_i)$, thus taking the difference between hopping away and hopping back probabilities would result in zero probability of staying on the same state – a problem present in Tully's original definition of hopping probabilities as well. Normalizing a combined "hopping-away-minus-hopping-back" probability is also unclear, with further problems when discussing cumulative hopping probabilities.

Instead, the following should be the criteria for hopping:

Looping from smallest j to largest,

- If $\left(\sum_s^j g^{+}_{ks}(t_i) > \xi\right)$

- If ($j = k$), no hop, exit loop.
- If ($j \neq k$),
    - If $\left(g^+_{kj}(t_i) > g^-_{kj}(t_i)\right)$, Try hop to state j on time step (t_{i+1}), exit loop.
    - Else, no hop, exit loop.

Note the absence of a max function since the probabilities are now always positive, the possibility of fulfilling the first inequality but not the second.

Using the propagator to determine hopping probabilities also locks the time interval that the probability is for, unlike Tully's probability which is determined for an instantaneous rate that simply needs to be present in the time interval. To be consistent, the propagators used have to be:

$$\mathbf{U}(t_i, t_{i-1}) = \mathbf{U}(t_i)\mathbf{U}(t_{i-1})$$

(following their definition $\mathbf{U}(t_i) = \exp(-i\mathbf{H}(t_i)\frac{\Delta t}{2})$ above), matching the time interval $[t_{i-1}, t_i)$.

The resulting algorithm is most simply written with the velocity verlet propagation of classical dynamics, and should be trivial to switch to leapfrog verlet.

3.3.1 Normalized Fewest switches surface hopping algorithm

A possible algorithm is as follows. From the previous step, $\mathbf{x}(t_i)$, $\mathbf{v}\,(t_{i-1})$, $\mathbf{F}_k(t_{i-1})$, and $\mathbf{U}_k(t_{i-1})$ are provided. Propagating the quantum density is not necessary, but is included, in which case $\boldsymbol{\rho}(t_{i-1})$ is needed as well.

For the first time step, $\mathbf{v}(t_0)$ is provided and $\mathbf{U}_k(t_{-1})$ is identity.

1. Calculate BO electronic structure properties.

$$\mathbf{V}(t_i) = \mathbf{V}(\mathbf{x}(t_i)), \mathbf{F}(t_i) = \mathbf{F}(\mathbf{x}(t_i)), \mathbf{d}(t_i) = \mathbf{d}(\mathbf{x}(t_i)) \tag{3.5}$$

2. Propagate classical nuclei velocity to current time step, to obtain old velocity corresponding to state k

$$\mathbf{v}_k\left(t_i\right) = \mathbf{v}\left(t_{i-1}\right) + \frac{1}{2\mathbf{m}}(\mathbf{F}_k(t_{i-1}) + \mathbf{F}_k(t_i))\Delta t \tag{3.6}$$

3. Construct model Hamiltonian and propagators with relevant classical nuclei velocity. Propagate quantum density using propagator. Reuse propagator from previous time step.

$$\mathbf{H}_k(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_k(t_i) \cdot \mathbf{d}(t_i) \tag{3.7}$$

$$\mathbf{U}_k(t_i) = \exp(-i\mathbf{H}_k(t_i)\frac{\Delta t}{2}) \tag{3.8}$$

$$\mathbf{U}(t_i, t_{i-1}) = \mathbf{U}_k(t_i)\mathbf{U}_k(t_{i-1}) \tag{3.9}$$

$$\boldsymbol{\rho}(t_i) = \mathbf{U}(t_i, t_{i-1})\, \boldsymbol{\rho}(t_{i-1})\, \mathbf{U}^{\dagger}(t_i, t_{i-1}) \tag{3.10}$$

4. Compute normalized Tully rates and hopping probabilities

$$g^{+}_{kj}(t_i) = \frac{U_{jk}(t_i, t_{i-1})U^{\dagger}_{kj}(t_i, t_{i-1})}{\sum_s U_{sk}(t_i, t_{i-1})U^{\dagger}_{ks}(t_i, t_{i-1})} \tag{3.11}$$

5. Decisions to attempt hop:

   - Looping from smallest l to largest: If $(\sum_s^l g^{+}_{ks}(t_i) > \xi)$,
     - If $(l = k)$, no hop, exit loop.

- If ($l \neq k$),

$$g_{kl}^{-}(t_i) = \frac{U_{kl}(t_i, t_{i-1}) U_{lk}^{\dagger}(t_i, t_{i-1})}{\sum_s U_{sl}(t_i, t_{i-1}) U_{ls}^{\dagger}(t_i, t_{i-1})} \tag{3.12}$$

  - If ($g_{kl}^{+}(t_i) > g_{kl}^{-}(t_i)$), attempt to hop to state l, exit loop.
  - Else, no hop, exit loop.

- If ($k = 1$ and $\epsilon_1 <$ thres), force a hop to state 0 and reorder density.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10}, \rho_{0s} \Leftrightarrow \rho_{1s}, \rho_{s0} \Leftrightarrow \rho_{s1} \tag{3.13}$$

- If attempting to hop, determine velocity rescale factor γ by solving

$$\frac{1}{2\mathbf{m}} \mathbf{d}_{lk}(t_i) \cdot \mathbf{d}_{lk}(t_i) \gamma^2 + \mathbf{v}_k(t_i) \cdot \mathbf{d}_{lk}(t_i) \gamma + \epsilon_l(t_i) - \epsilon_k(t_i) = 0 \tag{3.14}$$

- Decision to hop. The new active state is n
  - If $\gamma_{\pm}$ is real, select γ_{min} with the smallest magnitude and use it to rescale velocities, obtain a new velocity corresponding to state l: active state from t_i will be state l

$$\mathbf{v}_n(t_i) = \mathbf{v}_k(t_i) + \gamma_{\text{min}} \frac{\mathbf{d}_{lk}(t_i)}{\mathbf{m}}, n = l \tag{3.15}$$

  - If $\gamma_{\pm}$ is complex, hop is forbidden, active state remains k. No hop occurs. The same applies if no hop is attempted

$$\mathbf{v}_n(t_i) = \mathbf{v}_k(t_i), n = k \tag{3.16}$$

6. Propagate classical positions by a full step. Save forces for next time step.

$$\mathbf{x}(t_{i+1}) = \mathbf{x}(t_i) + \mathbf{v}_n\,(t_i)\,\Delta t + \frac{1}{2\mathbf{m}}\mathbf{F}_n(t_i)(\Delta t)^2 \tag{3.17}$$

7. Construct model Hamiltonian and propagators with relevant classical nuclei velocity, for next time step

$$\mathbf{H}_n(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_n(t_i)\cdot\mathbf{d}(t_i) \tag{3.18}$$

$$\mathbf{U}_n(t_i) = \exp(-i\mathbf{H}_n(t_i)\frac{\Delta t}{2}) \tag{3.19}$$

3.4 Improved switching

The more devastating problem with the simulations is total energy discontinuity. This problem arises from the sudden switching - energy rescaling and sudden changes in the forces on the nuclei.

There is however little choice but to do instantaneous rescaling of the nuclear velocities, as the electronic energy gaps are well defined only at fixed points in time. However, instead of changing immediately to a different state with very different forces, an Erenfest step can be used, where the system experiences an averaged force for the next full step. The energy gap to be rescaled would also be smaller. These will reduce the discontinuities, at the cost of hops taking two steps to complete.

There is a possibility of the hop to the averaged state being allowed, but the next hop back to one of the normal states being forbidden. As long as the time steps are small enough, the time spent on the averaged state can be interpreted like a forbidden hop, with the active state being reverted to the original active state. This treatment is consistent whether the hop is transferring energy from the nuclei to the electrons or vice versa.

3.4.1 Finite time step Tully surface hopping, with normalized probabilities and Ehrenfest step

A possible algorithm is as follows. From the previous step, $\mathbf{x}(t_i)$, $\mathbf{v}(t_{i-1})$, and $\mathbf{F}(t_{i-1})$ are provided, as well as information if a hop is being continued. The algorithm is not significantly affected by propagating on the Ehrenfest state. With the possibility of the mixed Ehrenfest state, the label *active* is used to generalize being on one adiabatic state or being on the Ehrenfest state. The most significant change is the propagator used for the hopping probability. A successful hop begins at time t_i based on past propagation and is completed at time t_{i+1}, which can be thought of as the time interval $[t_{i-1}, t_{i+1})$. It does appear that there is some double counting of time intervals for hopping, but it is necessary to ensure that every time step is treated equally and the fewest switches criteria should prevent rapid oscillations. The alternative to prevent time intervals for the hopping probability from overlapping is to skip the hopping probability calculation at every other time step, which would treat odd and even time steps inconsistently.

Superficially, it seems that this algorithm with hops to Ehrenfest states does require halving the time step size for direct comparison to previous algorithms, thus treating all time steps equivalently in this algorithm may be a flaw.

For the first time step, $\mathbf{v}(t_0)$ is provided.

1. Calculate BO electronic structure properties.

$$\mathbf{V}(t_i) = \mathbf{V}(\mathbf{x}(t_i)), \mathbf{F}(t_i) = \mathbf{F}(\mathbf{x}(t_i)), \mathbf{d}(t_i) = \mathbf{d}(\mathbf{x}(t_i)) \tag{3.20}$$

2. Propagate classical nuclei velocity to current time step, to obtain old velocity corre-

sponding to active state.

$$\mathbf{v}_{active}\left(t_i\right) = \mathbf{v}\left(t_{i-1}\right) + \frac{1}{2\mathbf{m}}(\mathbf{F}_{active}(t_{i-1}) + \mathbf{F}_{active}(t_i))\Delta t \tag{3.21}$$

3. If on Ehrenfest state averaged between k and l, attempt to complete hop off Ehrenfest state to adiabatic active state l:

   - determine velocity rescale factor γ_1 by solving

   $$\frac{1}{2\mathbf{m}}\mathbf{d}_{lk}(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma_1^2 + \mathbf{v}_{active}(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma_1 + \frac{\epsilon_l(t_i) - \epsilon_k(t_i)}{2} = 0 \tag{3.22}$$

   - determine velocity rescale factor γ_2 by solving

   $$\frac{1}{2\mathbf{m}}\mathbf{d}_{lk}(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma_2^2 + \mathbf{v}_{active}(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma_1 + \frac{\epsilon_k(t_i) - \epsilon_l(t_i)}{2} = 0 \tag{3.23}$$

   - Decision to hop. The new active state is n

     - If $\gamma_{1,\pm}$ is real, select $\gamma_{1,min}$ with the smallest magnitude and use it to rescale velocities, to obtain a new velocity corresponding to state l: active state from t_i will be state l

     $$\mathbf{v}_n(t_i) = \mathbf{v}_{active}(t_i) + \gamma_{1,\min}\frac{\mathbf{d}_{lk}(t_i)}{\mathbf{m}}, n = l \tag{3.24}$$

     - If $\gamma_{1,\pm}$ is complex, $\gamma_{2,\pm}$ is real, hop is reversed, active state returns to k.

     $$\mathbf{v}_n(t_i) = \mathbf{v}_{active}(t_i) + \gamma_{2,\min}\frac{\mathbf{d}_{lk}(t_i)}{\mathbf{m}}, n = k \tag{3.25}$$

4. Construct model Hamiltonian and propagators with relevant classical nuclei velocity.

$$\mathbf{H}_n(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_n(t_i) \cdot \mathbf{d}(t_i) \tag{3.26}$$

$$\mathbf{U}(t_i) = \exp(-2i\mathbf{H}_n(t_i)\Delta t) \tag{3.27}$$

5. Compute normalized Tully rates and hopping probabilities

$$g^+_{nj}(t_i) = \frac{U_{jn}(t_i)U^\dagger_{nj}(t_i)}{\sum_s U_{sn}(t_i)U^\dagger_{ns}(t_i)} \tag{3.28}$$

6. Decisions to attempt hop (can be done for all i or only for odd i):

- Looping from smallest q to largest: If $(\sum_s^q g^+_{ns}(t_i) > \xi)$,
  - If $(n = q)$, no hop, exit loop.
  - If $(n \neq q)$,

    $$g^-_{nq}(t_i) = \frac{U_{nq}(t_i)U^\dagger_{qn}(t_i)}{\sum_s U_{sq}(t_i)U^\dagger_{qs}(t_i)} \tag{3.29}$$

    - If $(g^+_{nq}(t_i) \geq g^-_{nq}(t_i))$ then attempt to hop to state q, exit loop.
    - Else, no hop, exit loop.
- If ($n = 1$ and $\epsilon_1 <$ thres), force a hop to state 0 and reorder density.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10}, \rho_{0s} \Leftrightarrow \rho_{1s}, \rho_{s0} \Leftrightarrow \rho_{s1} \tag{3.30}$$

- If attempting to hop, determine velocity rescale factor γ_3 by solving

$$\frac{1}{2\mathbf{m}}\mathbf{d}_{qn}(t_i) \cdot \mathbf{d}_{qn}(t_i)\gamma_3^2 + \mathbf{v}_n(t_i) \cdot \mathbf{d}_{qn}(t_i)\gamma_3 + \frac{\epsilon_q(t_i) - \epsilon_n(t_i)}{2} = 0 \tag{3.31}$$

- Decision to hop. The new active state state is labeled *active*

- If $\gamma_{3,\pm}$ is real, select $\gamma_{3,min}$ with the smallest magnitude and use it to rescale velocities, to obtain a new velocity corresponding to Ehrenfest state mixed between q and n: active state for t_i will be averaged state between n and q

$$\mathbf{v}_{active}(t_i) = \mathbf{v}_n(t_i) + \gamma_{3,\min} \frac{\mathbf{d}_{qn}(t_i)}{\mathbf{m}} \tag{3.32}$$

$$\mathbf{F}_{active}(t_i) = \frac{1}{2}(\mathbf{F}_n(t_i) + \mathbf{F}_q(t_i)) \tag{3.33}$$

- If $\gamma_{3,\pm}$ is complex, hop is forbidden, active state remains n. No hop occurs. The same applies if no hop is attempted

$$\mathbf{v}_{active}(t_i) = \mathbf{v}_n(t_i) \tag{3.34}$$

$$\mathbf{F}_{active}(t_i) = \mathbf{F}_n(t_i) \tag{3.35}$$

7. Propagate classical positions by a full step. Save forces for next step

$$\mathbf{x}(t_{i+1}) = \mathbf{x}(t_i) + \mathbf{v}_{active}\,(t_i)\,\Delta t + \frac{1}{2\mathbf{m}}\mathbf{F}_{active}(t_i)(\Delta t)^2 \tag{3.36}$$

3.5 Conclusions

Algorithms for Tully Surface Hopping that accounts for the finite time step interval has been proposed. The first algorithm solves the problem of normalizing Tully Surface Hopping probabilities, and the second could reduce the energy discontinuity from sudden-switching in the original algorithm.

Chapter 4

Characterizing UV-vis spectra of lanthanide complexes with different substituted cyclopentadiene ligands

Figure for experimental spectra taken from the following paper, Reprinted (adapted) with permission from *Structural, Spectroscopic, and Theoretical Comparison of Traditional vs Recently Discovered* Ln^{2+} *Ions in the* $[K(2.2.2-cryptand)][(C_5H_4SiMe_3)_3Ln]$ *Complexes: The Variable Nature of* Dy^{2+} *and* Nd^{2+} Megan E. Fieser, Matthew R. MacDonald, Brandon T. Krull, Jefferson E. Bates, Joseph W. Ziller, Filipp Furche, and William J. Evans **Journal of the American Chemical Society** 2015 137 (1), 369-382 DOI: 10.1021/ja510831n . Copyright 2015 American Chemical Society.

This chapter contains verbatim excerpts from the following paper. Reprinted (adapted) with permission from *Tetramethylcyclopentadienyl Ligands Allow Isolation of Ln(II) Ions across the Lanthanide Series in* $[K(2.2.2-cryptand)][(C_5Me_4H)_3Ln]$ *Complexes*, Tener F. Jenkins, David H. Woen, Luke N. Mohanam, Joseph W. Ziller, Filipp Furche, and William

J. Evans **Organometallics** 2018, 37 (21), 3863-3873 DOI: 10.1021/acs.organomet.8b00557 . Copyright 2018 American Chemical Society.

The material in this chapter is based upon work supported by the National Science Foundation under CHE-1800431.

For visualization of orbitals, self consistent Kohn-Sham orbitals are used with a contour value of 0.04. For visualizing predicted spectra, individual excitations are broadened with Gaussians with a root mean square width of 0.15 eV [32], and vertical scaling of factor 0.25. A uniform empirical blue shift of 0.15 eV was also used to align the predicted spectra with the experiment.

All calculations were performed with the TURBOMOLE 7.2 package

(https://www.turbomole.org/).

Analysis was aided by VMD and Avogadro visualization tools, available from

https://www.ks.uiuc.edu/Research/vmd/

and

https://avogadro.cc/

respectively.

4.1 Chapter Summary

Although previous studies of the stabilization of Ln(II) ions across the lanthanide series have relied on Me_3Si-substituted cyclopentadienyl ligands, we now find surprisingly that these ions can also exist surrounded by three tetramethylcyclopentadienyl ligands. Reduction of the

$4f^n$ Ln(III) complexes, $Cp_3^{tet}Ln$ ($Cp^{tet} = C_5Me_4H$) using potassium graphite in the presence of 2.2.2-cryptand (crypt) produces the Ln(II) complexes, $[K(crypt)][Cp_3^{tet}Ln]$ for Ln = La, Ce, Pr, Nd, Sm, Gd, Tb, Dy, all of which were characterized by X-ray crystallography. These complexes display intense absorptions in the UV-visible-near IR region that are red-shifted compared to those of previously characterized $(Cp'_3Ln)^{1-}$ complexes ($Cp' = C_5H_4SiMe_3$).

The reference ground state electronic structure was determined with *ab-initio* Density Functional Theory and molecular geometry optimization, with comparative calculations with different parameters. Linear response calculations from the reference state, in the framework of Time-Dependent Density Functional Theory, were used to predict the electronic excitations in the UV-visible region of the spectra, allowing for assignment of the experimental absorption peaks. Basis set augmentation was required and justified.

4.2 Background and experimental ground state characterization

In recent years, the range of oxidation states available to the rare-earth metals in crystallographically-characterizable molecular complexes available for reactivity in solution has greatly expanded.[58, 59] Up until 2001, it was thought that only six lanthanides could form crystallographically-characterizable molecular complexes of Ln(II) ions in solution: Eu, Yb, Sm, Tm, Dy, and Nd. [60, 61, 62, 63] These complexes could be made by reduction of $4f^n$ Ln(III) precursors and formed Ln(II) ions with $4f^{n+1}$ electron configurations as expected. However, it is now known that yttrium and all of the lanthanides (except Pm which was not studied due to its radioactivity) can form isolable molecular complexes of Ln(II) ions if reductions are done in the proper coordination environment.[58, 59] Specifically, reduction of tris(cyclopentadienyl) complexes with silyl-substituted ligands $C_5H_3(SiMe_3)_2$ (Cp") [64, 65] and $C_5H_4SiMe_3$ (Cp')[66, 67, 68, 69] provided access to Ln(II) ions across the series.[58, 59] This was also extended to the actinides, Th,[70] U,[71, 72] Pu,[73] and Np.[74] Examples of complexes of new Ln(II) ions are also known with $C_5H_3(CMe_3)_2$ (Cp^{tt})[75, 76] and $C_5H_2(CMe_3)_3$ (Cp^{ttt})[77, 78] and the tris(aryloxide) mesitylene ligand, $[(^{Ad,Me}ArO)_3mes]^{3-}$.[79, 80] Extensive crystallographic, spectroscopic, magnetic, and density functional theory (DFT) studies showed that the new Ln(II) ions in the tris(cyclopentadienyl) environments adopted $4f^n5d^1$ electron configurations.[58, 64, 65, 66, 67, 68, 69, 80, 81, 82, 83] This could be rationalized by the crystal field splitting of a tris(cyclopentadienyl) coordination environment, which puts a $5d_z^2$ orbital comparable in energy to the $4f$ orbitals.[84, 85, 86, 87, 88, 89]

A comparison of the rare-earth metals in the [K(crypt)][Cp''_3Ln], series (crypt = 2.2.2-cryptand) revealed that Ln(II) ions could be grouped into three categories.[67] Sm, Eu, Tm, and Yb form complexes of traditional Ln(II) ions with $4f^{n+1}$ electron configurations. For Ln = Nd and Dy, the [K(crypt)][Cp'_3Ln] complexes have $4f^n5d^1$ configurations, but in other ligand environments the metals form $4f^{n+1}$ Ln(II) ions. These are designated as configurational crossover ions. Nd(II) and Pr(II) are known to be configurational crossover ions in the solid state iodides, LnI_2.[90, 91] The third category contains the rest of the lanthanide metals which have $4f^n5d^1$ configurations and Y(II) which is a $4d^1$ ion.

The recent isolation of +2 ions in complexes of NR_2 (R = $SiMe_3$) ligands, suggested that strongly donating ligands could also provide these new ions. Specifically, amide ligands were used to isolate the first crystallographically-characterizable complex of a +2 ion of the smallest rare-earth metal, scandium, as well as the lanthanides, Ln = Nd, Gd, Tb, Dy, Ho, and Er, in the compounds [M(chelate)][$Ln(NR_2)_3$] (M = K, Rb; chelate = 18-crown-6 and crypt; R = $SiMe_3$).[92, 93] Consequently, we have examined reduction of $Cp_3^{tet}Ln$ complexes (Cp^{tet} = C_5Me_4H). These complexes were not examined earlier [82] because it was assumed that the Cp^{tet} ligand was too electron donating to form stable complexes. We report here that the $(Cp_3^{tet})^{3-}$ ligand set provides an entire new *series* of Ln(II) complexes that allows evaluation of the three categories of Ln(II) complexes as a function of metal.

4.3 Experimental ground state molecular geometry

Metrical parameters on the [K(crypt)][Cp_3^{tet}Ln], 2-Ln, series are presented in Table 4.1, along with data on the analogous [K(crypt)][Cp'$_3$Ln], 3-Ln, [66, 67, 68, 69] and [K(crypt)][Cp"$_3$Ln][64, 65] complexes when available. The Ln-(C_5Me_4H ring centroid) distances (Ln-Cnt) for six of the eight 2-Ln complexes with Ln = La, Ce, Pr, Nd, Gd, and Tb are consistent with Ln ionic radii, decreasing with increasing atomic number, following the lanthanide contraction. For each of these six metals, the Ln-Cnt distance also decreases in the order of the size of the ligands: Cp^{tet} >> Cp" > Cp'. This suggests that the $(Cp_3^{tet})^{3-}$ environment occupies more space than (Cp"$_3$)$^{3-}$ which is surprising given that (Cp")$^{1-}$ was investigated as a ligand for being sterically similar to $(C_5Me_5)^{1-}$.[94]

The differences in Ln-Cnt distances between the Ln(III) $(C_5R_5)_3$Ln precursor and the reduced Ln(II) product for La, Ce, Pr, Nd, Gd, and Tb are 0.045-0.058 Å for the Cp^{tet} series versus 0.018-0.022 Å for the Cp" complexes and 0.026-0.031 Å for the Cp' compounds. All of these differences are much smaller than the 0.1-0.2 Å differences found for $4f^{n+1}$ Ln(II) versus $4f^n$ Ln(III) complexes of the traditional Ln(II) ions of Eu, Yb, Sm, and Tm. This is structural evidence consistent with $4f^n5d^1$ configurations for these six metals based on previous structural, spectroscopic, and DFT analysis of the 3-Ln complexes. [66, 67, 68, 69]

4.4 Experimental UV-visible spectra

Another notable difference between 2-Ln, and 3-Ln involves the UV-vis spectra. The absorptions with the largest λ_{max} values are at lower energies compared to the most intense

	Ln-Cnt Range (Å)	Ln-Cnt Average (Å)	Δ[Ln(II) vs Ln(III)]*a*	Ref.
(Cp$_3^{tet}$La)$^{1-}$	2.626 – 2.642	2.633	0.058	[95]
(Cp”$_3$La)$^{1-}$	2.606 – 2.642	2.620	0.018	[64, 96]
(Cp’$_3$La)$^{1-}$	2.581 – 2.595	2.586	0.026	[67, 97]
(Cp$_3^{tet}$Ce)$^{1-}$	2.594 – 2.612	2.603	0.051	[98]
(Cp”$_3$Ce)$^{1-}$	2.574 – 2.609	2.587	0.022	[65, 99]
(Cp’$_3$Ce)$^{1-}$	2.553 – 2.567	2.558	0.029	[67, 99]
(Cp$_3^{tet}$Pr)$^{1-}$	2.572 – 2.583	2.578	0.046	[98]
(Cp”$_3$Pr)$^{1-}$	2.552 – 2.588	2.566	*b*	
(Cp’$_3$Pr)$^{1-}$	2.530 – 2.544	2.535	0.026	[66, 100]
(Cp$_3^{tet}$Nd)$^{1-}$	2.555 – 2.568	2.563	0.045	[101]
(Cp”$_3$Nd)$^{1-}$	2.530 – 2.559	2.544	0.019	[65, 96]
(Cp’$_3$Nd)$^{1-}$	2.514 – 2.528	2.519	0.031	[67, 100]
(Cp$_3^{tet}$Sm)$^{1-}$	2.623 – 2.640	2.630	0.147	[95]
(Cp’$_3$Sm)$^{1-}$	2.603 – 2.615	2.608	0.148	[67]
(Cp$_3^{tet}$Gd)$^{1-}$	2.511 – 2.519	2.516	0.047	
(Cp’$_3$Gd)$^{1-}$	2.463 – 2.475	2.468	0.031	[66]
(Cp$_3^{tet}$Tb)$^{1-}$	2.498 – 2.505	2.502	0.054	[95]
(Cp’$_3$Tb)$^{1-}$	2.448 – 2.461	2.454	0.032	[66]
(Cp$_3^{tet}$Dy)$^{1-}$	2.543 – 2.543	2.543	0.099	
(Cp’$_3$Dy)$^{1-}$	2.434 – 2.450	2.443	0.036	[67]

Table 4.1: Comparison of Ln–(cyclopentadienyl ring centroid) distances (Ln–Cnt) for [K(crypt)][CpX$_3$Ln] with CpX = Cp^{tet}, Cp’, and Cp”.
*a*Δ[Ln(II) vs Ln(III)] = the difference in Ln-Cnt distances of [CpX$_3$Ln^{II}]$^{1-}$ vs CpX$_3$Ln^{III}. The references are to the structures of the CpX$_3$Ln^{III} complexes.
b The structure of Cp''_3Pr has not been reported for comparison. [65]

absorptions of 3-Ln for Ln = La, Ce, Pr, Nd, Gd, and Tb.

The UV-vis spectra of 2-Ln are shown in Figure 4.1 and the absorbance maxima and extinction coefficients are compared with those of 3-Ln in Table 4.2. The spectra of 2-Ln for Ln = La, Ce, Pr, Nd, Gd, and Tb show the strongest absorptions in the near-infrared region with λ_{max} values in the range of 745-874 nm with a molar extinction coefficient, ϵ, of 1700-5600 $M^{-1}cm^{-1}$. In comparison, the previously reported spectra of 3-Ln complexes (Figure 4.2) of these metals have the largest absorptions from 420-635 nm in the visible region with ϵ = 4400-6500 $M^{-1}cm^{-1}$. [58, 59, 66, 67, 68, 69]

In contrast to the UV-vis spectra of La, Ce, Pr, Nd, Gd, and Tb, the spectra of 2-Dy and 2-Sm have weaker absorptions. These are shown with a different scale in Figure 4.1 (bottom figure). Less intense absorptions were previously observed for the 3-Ln complexes for the metals with $4f^{n+1}$ electron configurations, i.e. Eu, Yb, Sm, and Tm.[67] Hence, the λ_{max} = 466 nm with $\epsilon = 203\ M^{-1}cm^{-1}$ for 2-Sm is consistent with a $4f^6$ electron configuration for this complex. The weaker absorption for 2-Dy with λ_{max} = 766 nm and $\epsilon = 200\ M^{-1}cm^{-1}$ is also consistent with a $4f^{n+1}$ configuration and this matches the structural data above which suggested that 2-Dy had a $4f^{10}$ configuration.

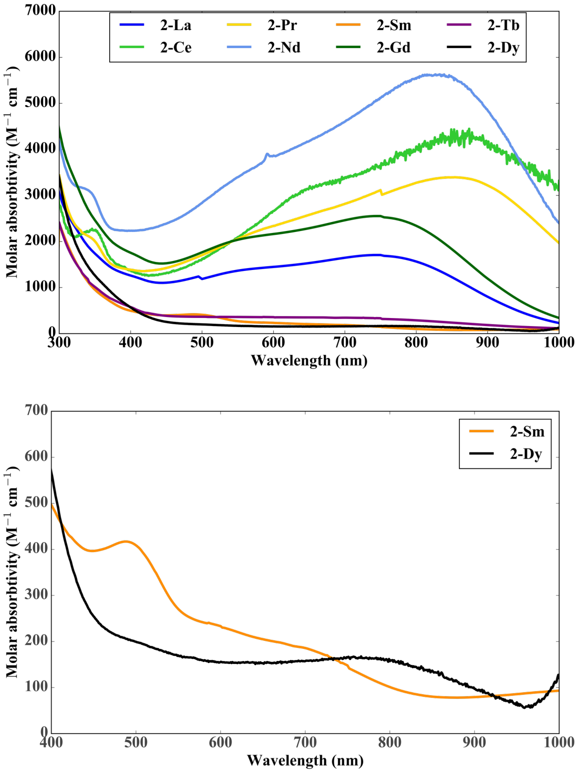

Figure 4.1: Experimental UV-vis absorption spectra of 2-Ln

4.5 Calculation of reference ground state

The structures of the $Cp_3^{tet}Ln$ precursors and the $(Cp_3^{tet}Ln)^{1-}$ anions in the 2-Ln complexes were optimized with density functional theory (DFT) using the Tao–Perdew–Staroverov–Scuseria

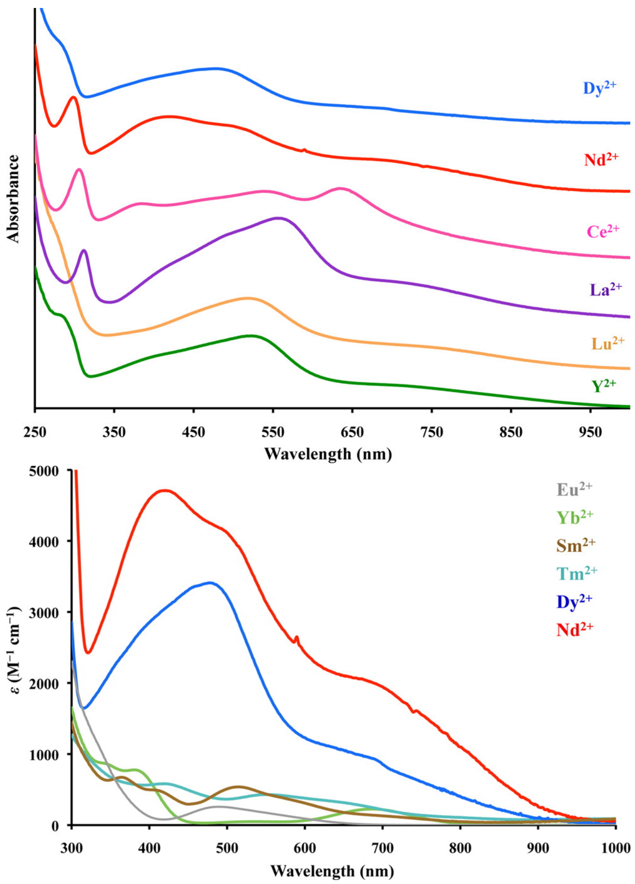

Figure 4.2: Experimental UV-vis absorption spectra of 3-Ln [67]

meta-generalized gradient approximation hybrid (TPSSh) density functional along with split valence basis sets [23] for the ligands and quasi-relativistic f-in-core effective core potentials (ECPs) [102] for 2-La, 3-La, 2-Ce, 3-Ce, 2-Sm, 3-Sm, 2-Gd and 3-Gd; f-in-core electron configurations are denoted by brackets in the following discussion. This methodology was previously established for the 3-Ln series [66] and further validated by f-out-of-core calculations [103] for 2-La, 3-La, 2-Ce, 3-Ce, 3-Pr, 2-Sm, 3-Sm, 2-Gd and 3-Gd.

Initial structures were generated from X-ray crystal structure of 2-Ln and 3-Ln (Ln = La, Ce, Pr, Nd, Gd, Tb). Solvent effects were included using the COSMO continuum solvent model [104], with the following parameters to model the THF solvent environment: dielectric constant = 7.520, refractive index = 1.405 [105]. Due to the longer metal-ligand distances in the 2-Ln complexes, it was necessary to increase the lanthanide metal atom radii used

	$\lambda_{max}(nm)$	$\epsilon(M^{-1}cm^{-1})$
2-La	745	1700
3-La[67]	554	6500
2-Ce	874	4100
3-Ce[67]	635	4700
2-Pr	854	4500
3-Pr[66]	518	4500
2-Nd	833	5600
3-Nd[67]	420	4700
2-Sm	466	200
3-Sm[67]	360	700
2-Gd	745	2500
3-Gd[66]	430	4400
2-Tb	784	650
3-Tb[66]	464	4800
2-Dy	766	200
3-Dy[67]	483	3400

Table 4.2: Comparison of UV-vis absorption maxima of [K(crypt)][Cp_3^{tet}Ln], 2-Ln, and [K(crypt)][Cp'$_3$Ln], 3-Ln.

to construct the cavity to 2.584 angstroms (from the default 2.223 angstroms for all lanthanides), to obtain a continuous cavity. The default solvent radii (1.30 angstroms) was used. C_1 molecular symmetry was used for all calculations.

Density functional theory (DFT) structural optimizations of 2-Ln and 3-Ln (Ln= La, Ce, Pr, Nd, Sm, Gd, Tb, Dy) were performed using the hybrid meta-generalized gradient approximation (meta-GGA) functional, TPSSh [106]. Grimme's dispersion corrections were used[107]; the resolution of identity approximation (RI-J) was employed in all calculations.[51] DFT quadrature grids were of size 4 or larger in conjunction with tight energy convergence threshold of 10^{-8} Hartrees and one-electron density convergence threshold of 10^{-8}. [16] The geometry convergence threshold was 10^{-5} a.u. For atoms on the ligands, double-ζ quality split valence basis sets with polarization functions on non-hydrogen atoms [def2-SV(P)] were used [23]. For the Lanthanide atoms, large f-in-core quasi-relativistic Stuttgart-Cologne effective core potentials (ECPs) with corresponding triple-ζ quality basis sets [SCecp-mwb] were used.

[102, 108, 109, 110] The metal-ligand distance is reported in table 4.4.

For validation of the f-in-core results, DFT f-out-of-core structural optimization of 2-Ln were attempted, substituting the f-in-core basis set and ECP on the metal with a double-ζ quality split valence basis set and ECP with polarization functions on non-hydrogen atoms [def2-SV(P)] [103, 23] in the procedure above. All converged results for Ln = La, Ce, Pr, Nd, Sm and Gd are reported, where some calculations converge to different electronic structures. A comparison of the number of core electrons treated by the ECP is reported in Table 4.3.

	f-out-of-core def2-SV(P)	f-in-core[$4f^n$] SCecp-mwb	f-in-core[$4f^{n+1}$] SCecp-mwb
La	46	46	
Ce	28	47	
Pr	28	48	49
Nd	28	49	50
Sm	28	51	52
Gd	28	53	
Tb		54	55
Dy		55	56

Table 4.3: Comparison of number of implicit electrons in standard ECPs
This information is well known and found in reference [102, 108, 109, 110, 23] , and is restated for clarity.

The f-out-of-core DFT calculations produced states with HOMOs showing mixed f- and d-character for 2-Nd and 2-Pr, with greater f-character. Even though the agreement with the experimental structures is somewhat worse than with the [$4f^n$]$5d^1$ calculations, the differences are hardly significant. Thus, the f-out-of-core calculations appear to slightly overestimate the stability of the $4f^{n+1}$ configuration compared to experiment; this is consistent with prior observations for the Cp' complexes.[67] Given the inherent inaccuracy of DFT implementations for such multi-configurational states, this is not surprising. The f-in-core calculations have less f-functions on the metal, avoiding the overestimation with the exclusion of significant f-character in the MOs calculated.

For the La, Ce, Pr, Nd, Gd, and Tb anions in 2-Ln, the calculated average Ln-(Cp ring centroid) distances are within 0.001-0.02 Å of the experimental values. Moreover, for these six metals, the computed change in Ln-Cnt distances between the Ln(III) precursor Cp_3^{tet}Ln and the reduced Ln(II) product, 2-Ln, matches the X-ray data within 0.02 Å when $[4f^n]5d^1$ configurations for the anions are assumed with the f-in-core effective core potentials (Table 4.4). The highest occupied molecular orbital of 2-La is shown in Figure 4.3. The d_{z^2} nature of this HOMO matches those found for 3-Ln.

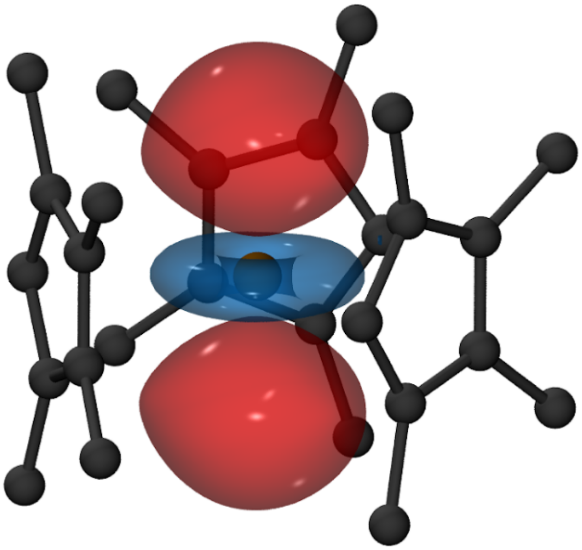

Figure 4.3: Singly occupied αHOMO of 2-La, showing $5d_{z^2}$ character

Since Nd(II) is a configurational crossover ion that has been observed to access both $4f^4$ and $4f^3 5d^1$ electron configurations, calculations on $(Cp_3^{tet}Nd)^{1-}$ with a $[4f^4]$ electron configuration were also carried out. With this traditional electronic configuration, a 2.727 Å Nd-Cnt distance is predicted which is much longer than the experimentally observed value of 2.563 Å. Also, no appreciable absorptions are predicted in the visible spectrum, in contrast to the strong experimental absorbance. The f-out-of-core calculations converged to a predominantly $4f^4$ state, but with a Nd–centroid distance of 2.621 Å. Calculations with a $[4f^3]5d^1$ configurations gave a value of 2.580 Å, closer to the metrical parameters from the

X-ray crystal structure, and excitation spectra in qualitative agreement with experiment. These results provide support for the assignment of 2-Nd as a $4f^35d^1$ ion with a low-lying excited $4f^4$ configuration. In contrast to the results above, calculations on the traditional Sm(II) ion with a $[4f^6]$ electron configuration in $(\mathrm{Cp}_3^{tet}\mathrm{Sm})^{1-}$ led to a 2.696 Å, Sm-Cnt distance which is longer than the 2.630 Å observed distance. However, f-out-of-core calculations predicted a $4f^6$ configuration with a 2.610 Å Sm–centroid distance in close agreement with experiment. For the configurational crossover ion Dy(II), the calculated bond distance for neither the $[4f^10]$ configuration, 2.650 Å, nor the $[4f^9]5d^1$ configuration, 2.505 Å, matched the 2.543 Å observed value. However, the calculated differences in Ln-ring centroid distances of the Ln(III) $\mathrm{Cp}_3^{tet}\mathrm{Ln}$ and the Ln(II) $(\mathrm{Cp}_3^{tet}\mathrm{Ln})^{1-}$ for Sm and Dy, 0.200 and 0.199 Å, respectively, were much larger than those of the $4f^n5d^1$ ions described above. The experimentally observed Ln(II) vs Ln(III) differences for Sm and Dy, 0.147 and 0.099 Å, are not as large as the calculated differences, but the data support the presence of $4f^{n+1}$ configurations for the ions in $(\mathrm{Cp}_3^{tet}\mathrm{Sm})^{1-}$ and $(\mathrm{Cp}_3^{tet}\mathrm{Dy})^{1-}$ which is consistent with the UV-vis data. Time-dependent density functional theory (TDDFT) simulations are consistent with this view and qualitatively reproduce the much stronger visible absorption of the La, Ce, Pr, Nd, Gd, and Tb compounds compared to the Sm and Dy compounds, see Figure 4.4.

4.5.1 Predicted Δ[Ln(II) vs Ln(III)]

For further validation of the calculated geometry, DFT structural optimization of $[Ln(Cp^{tet})_3]$ (Ln= La, Ce, Pr, Nd, Sm, Gd, Tb, Dy), the oxidised form of 2-Ln, were carried out with identical setting to the structural optimization described above, to predict the change in metal-ligand distance on reduction, which is known experimentally(Table 4.1). Since these are f-in-core calculations, the ECP used for Sm and Dy differ between 2-Ln and $[\mathrm{Ln}(\mathrm{Cp}^{tet})_3]$ due to the different number of f-electrons, as reported in Table 4.3.

4.6 Predicted UV-vis spectra

For the time dependent density functional theory (TDDFT) calculation[10, 1] of compounds 2-Ln and 3-Ln (Ln= La, Ce, Pr, Nd, Gd, Tb), the lanthanide basis sets were augmented with an additional diffuse *p*-function by downward extrapolation [SCecp-mwb-d], with the exponents reported in table 4.7. A gauge invariant implementation of the TPSSh functional was used. [111] For the ligand atoms, augmented polarized split valence basis (def2-SVPD[112]) were employed. To simulate the visible absorption spectra, TDDFT calculations for the 25 lowest electronic excitations were run (Table C.1 through C.12, Figure 4.4 and Figure C.1 through C.6).

In addition, to simulate the visible absorption spectra of 2-Sm and 3-Sm, TDDFT calculations for the 25 lowest electronic excitations were run, using a gauge invariant implementation of the TPSSh functional [111] but without basis set augmentation. (Figure 4.5)

TDDFT studies indicate that the maximum absorbances for the two series do not arise from analogous transitions. Analysis of time-dependent density functional theory (TDDFT) simulations for the $(\mathrm{Cp}_3^{tet}\mathrm{Ln})^{1-}$ (Ln = La, Ce, Pr, Nd, Gd, and Tb) compounds reveals that in addition to the MLCT absorptions found previously in the 3-Ln series, there are strongly dipole-allowed Ln $5d \rightarrow 6p$ and $5d \rightarrow \pi^*$ transitions (Table C.1 through C.12) at lower energy. The three distinct bands in the computed visible spectra can be associated with transitions to $6p$ Rydberg orbitals of slightly different energies. The higher oscillator strengths for the $5d \rightarrow 6p$ transitions in the 2-Ln series can be rationalized by the larger metal-ligand distances in these compounds, which stabilize metal $6p$ Rydberg orbitals relative to the 3-Ln compounds. MLCT transitions are still present in 2-Ln (Ln = La, Ce, Pr, Nd, Gd, and Tb), but they are much weaker than these $5d \rightarrow 6p$ transitions. It should also be noted that $5d \rightarrow 6p$ absorptions are also present in 3-Ln, but are weaker in those complexes. For 2-Ln, the most intense absorptions arise from Ln $5d \rightarrow 6p$ transitions; for 3-Ln the largest absorptions are attributed to MLCT. While both transitions are present in both complexes, the MLCT transitions are more intense for 3-Ln and the metal-metal transitions are more intense for 2-Ln. It is likely that the longer metal-ligand distances in 2-Ln weaken the MLCT and enhance the $5d \rightarrow 6p$ transitions.

The difference in intensities of the two series could also occur due to the difference in the geometry of the complexes. Geometric effects may also be operative in the difference in intensities of the transitions in 2-Ln vs 3-Ln. The C_{3h} 2-Ln complexes are more symmetrical when compared to the 3-Ln species. In the 3-Ln series, two of the Me_3Si substituents are above the plane defined by the metal and the three Cp' ring centroids and one Me_3Si group is below. In the 2-Ln series, the arrangement of substituents has a different symmetry: the one unique ring carbon atom is the C-H unit rather than the unique ring substituted C-$SiMe_3$ moiety in 3-Ln. In 2-Ln, these unique C-H positions are more similar in the three rings and are in or near the plane of the three ring centroids. As a result, there are six methyl groups above the plane of the three Cp^{tet} ring centroids and six below. In addition to the changes in metal-ligand distance, the higher molecular symmetry of the 2-Ln complexes could be the cause of greater metal p-character in lower energy unoccupied MOs, which would have a stronger oscillator strength due to a large $5d \rightarrow 6p$ transition dipole moment.

For validation of the basis set augmentation, a TDDFT calculation was carried out on 2-La , using the f-out-of-core geometry and augmenting the basis set and ECP for all atoms with double-ζ quality split valence basis sets with polarization and diffuse functions [def2-SVPD][112]. This basis set and ECP is not available for the other lanthanides. The response calculation was carried out for the lowest 25 roots or excitations with the same functional, with a default convergence threshold of 10^{-5}.(Table C.13)

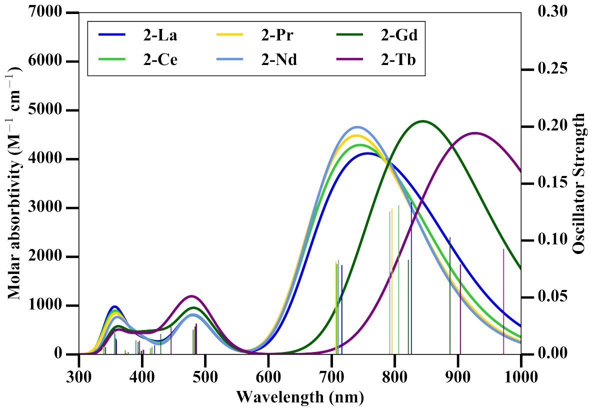

Figure 4.4: Predicted UV-vis absorption spectra for 2-Ln

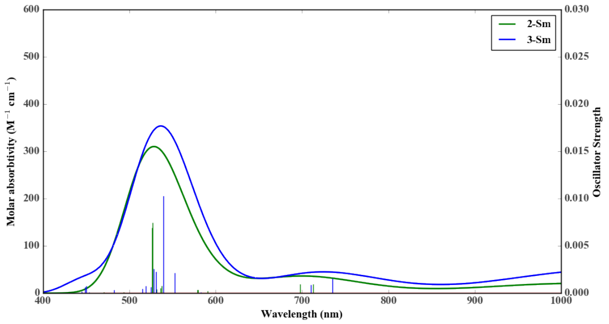

Figure 4.5: Predicted UV-vis absorption spectra for 2-Sm and 3-Sm

4.6.1 Verifying solvent non-participation

From comparing the calculated spectra with the experimental spectra, there is evidence of a relatively low energy, empty and diffuse molecular orbital with 6p_z character from the metal in these complexes. For 2-Ln, this empty orbital and the filled orbitals are orientated in space such that one or two THF solvent molecules may be involved in a dative bond with the metal with π back bonding – a possibility excluded by geometry optimization for 2-La (largest metal-ligand distance) and 2-Gd (most unpaired electrons) with explicit THF molecules, using the TPSSh functional and a larger triple-zeta quality basis sets with polarization and diffuse functions on all atoms[def2-TZVP]. [23] The closest approach of THF to the metal center is reported in table 4.8.

4.7 Conclusion

The isolation of $(\mathrm{Cp}_3^{tet}\mathrm{Ln})^{1-}$ complexes for Ln = La, Ce, Pr, Nd, Gd, and Tb indicates that the prior assumption that electron-rich cyclopentadienyl ligands cannot be used to isolate $4f^n5d^1$ Ln(II) complexes is incorrect. More generally, this suggests that the new mixed configuration Ln(II) ions are likely to be accessible with a variety of ligand systems. This will allow a broader investigation of both their physical properties and their reactivity. In addition, the results show the importance of the ligand in determining which configuration, $4f^{n+1}$ or $4f^n5d^1$, is adopted by Ln(II) ions in tris(cyclopentadienyl) coordination environments. The crossover point between the two configurations is variable depending on the ligands. Thus, judicious choice of the ligand environment may profoundly affect the chemical, optical, and magnetic properties of divalent lanthanides which can be significantly different for $4f^{n+1}$ vs $4f^n5d^1$ configurations.

			Ln-Cnt Range(Å)	Ln-Cnt Average(Å)	Difference from experiment (Å)
$(\mathrm{Cp}_3^{tet}\mathrm{La})^{1-}$	$[4f^n]5d^1$	(in core)	$2.632 - 2.637$	2.634	0.001
	$4f^n5d^1$	(out of core)	$2.665 - 2.668$	2.666	0.033
$(\mathrm{Cp}_3^{tet}\mathrm{Ce})^{1-}$	$[4f^n]5d^1$	(in core)	$2.615 - 2.619$	2.617	0.014
	$4f^n5d^1$	(out of core)	$2.586 - 2.598$	2.593	-0.01
$(\mathrm{Cp}_3^{tet}\mathrm{Pr})^{1-}$	$[4f^n]5d^1$	(in core)	$2.594 - 2.598$	2.596	0.018
	$[4f^{n+1}]$	(in core)	$2.741 - 2.745$	2.743	0.165
	$4f^{n+1}$	(out of core)	$2.598 - 2.609$	2.604	0.026
$(\mathrm{Cp}_3^{tet}\mathrm{Nd})^{1-}$	$[4f^n]5d^1$	(in core)	$2.578 - 2.581$	2.580	0.017
	$[4f^{n+1}]$	(in core)	$2.725 - 2.730$	2.727	0.164
	$4f^{n+1}$	(out of core)	$2.613 - 2.637$	2.621	0.058
$(\mathrm{Cp}_3^{tet}\mathrm{Sm})^{1-}$	$[4f^{n+1}]$	(in core)	$2.693 - 2.699$	2.696	0.066
	$4f^{n+1}$	(out of core)	$2.607 - 2.613$	2.610	-0.020
$(\mathrm{Cp}_3^{tet}\mathrm{Gd})^{1-}$	$[4f^n]5d^1$	(in core)	$2.521 - 2.528$	2.526	0.010
	$4f^n5d^1$	(out of core)	$2.587 - 2.589$	2.590	0.074
$(\mathrm{Cp}_3^{tet}\mathrm{Tb})^{1-}$	$[4f^n]5d^1$	(in core)	$2.507 - 2.517$	2.513	0.011
	$[4f^{n+1}]$	(in core)	$2.645 - 2.653$	2.649	0.147
$(\mathrm{Cp}_3^{tet}\mathrm{Dy})^{1-}$	$[4f^{n+1}]$	(in core)	$2.647 - 2.653$	2.650	0.107
	$[4f^n]5d^1$	(in core)	$2.503 - 2.507$	2.505	-0.038

Table 4.4: Comparison of calculated geometry with different parameters to experiment, for 2-Ln

			Ln-Cnt Range(Å)	Ln-Cnt Average(Å)	Difference from experiment (Å)
$(\text{Cp'}_3\text{La})^{1-}$	$[4f^n]5d^1$	(in core)	2.590 – 2.645	2.614	0.028
	$4f^n5d^1$	(out of core)	2.625 – 2.679	2.648	0.062
$(\text{Cp'}_3\text{Ce})^{1-}$	$[4f^n]5d^1$	(in core)	2.568 – 2.623	2.592	0.034
	$4f^n5d^1$	(out of core)	2.544 – 2.598	2.565	0.007
$(\text{Cp'}_3\text{Pr})^{1-}$	$[4f^n]5d^1$	(in core)	2.550 – 2.597	2.570	0.035
	$[4f^{n+1}]$	(in core)	2.741 – 2.745	2.743	0.165
	$4f^{n+1}$	(out of core)	2.625 – 2.679	2.647	0.112
$(\text{Cp'}_3\text{Nd})^{1-}$	$[4f^n]5d^1$	(in core)	2.531 – 2.582	2.553	0.034
	$4f^{n+1}$	(out of core)	2.590 – 2.654	2.617	0.098
$(\text{Cp'}_3\text{Sm})^{1-}$	$[4f^{n+1}]$	(in core)	2.686 – 2.735	2.703	0.095
	$4f^{n+1}$	(out of core)	2.589 – 2.635	2.605	−0.003
$(\text{Cp'}_3\text{Gd})^{1-}$	$[4f^n]5d^1$	(in core)	2.469 – 2.517	2.488	0.020
	$4f^n5d^1$	(out of core)	2.554 – 2.598	2.570	0.102
$(\text{Cp'}_3\text{Tb})^{1-}$	$[4f^n]5d^1$	(in core)	2.453 – 2.499	2.472	0.018
$(\text{Cp'}_3\text{Dy})^{1-}$	$[4f^n]5d^1$	(in core)	2.455 – 2.504	2.475	0.032

Table 4.5: Comparison of calculated geometry with different parameters to experiment, for 3-Ln

$[4f^n]$	Ln-Cnt Range (Å)	Ln-Cnt Average (Å)	Predicted $\Delta[Ln(II) vs Ln(III)]$(Å)	Difference in Δ from experiment (Å)
$(\text{Cp}_3^{tet}\text{La})$	2.576 – 2.577	2.576	0.058	0.000
$(\text{Cp}_3^{tet}\text{Ce})$	2.556 – 2.557	2.557	0.060	0.009
$(\text{Cp}_3^{tet}\text{Pr})$	2.538 – 2.539	2.538	0.058	0.012
$(\text{Cp}_3^{tet}\text{Nd})$	2.521 – 2.526	2.524	0.056	0.011
$(\text{Cp}_3^{tet}\text{Sm})$	2.494 – 2.497	2.496	0.200	0.053
$(\text{Cp}_3^{tet}\text{Gd})$	2.468 – 2.472	2.471	0.055	0.008
$(\text{Cp}_3^{tet}\text{Tb})$	2.494 – 2.497	2.496	0.017	−0.037
$(\text{Cp}_3^{tet}\text{Dy})$	2.449 – 2.453	2.451	0.199	0.100

Table 4.6: Comparison of Δ[Ln(II) vs Ln(III)] between calculation and experiment

	Exponent (a.u.)
La	$0.25199876548E-01$
Ce	$0.26032053387E-01$
Pr	$0.26626581360E-01$
Nd	$0.26680133595E-01$
Sm	$0.90125672267E-02$
Gd	$0.20297158144E-01$
Tb	$0.15975284823E-01$
Dy	$0.63206507004E-02$

Table 4.7: Exponents of the downward extrapolated p-function added to obtain the SCecp-mwb-d basis

printed by TURBOMOLE 7.2

	Metal-Solvent Hydrogen distance (Å)
2-La	$3.928, 4.083$
2-Gd	$4.147, 4.562$

Table 4.8: Explicit solvent's nearest hydrogen distance from the metal center

Chapter 5

Pedagogy Study for teaching intermolecular forces

This chapter is quoted verbatim from the following paper. Reprinted (adapted) with permission from *Intermolecular Forces Game: Using a Card Game to Engage Students in Reviewing Intermolecular Forces and Their Relationship to Boiling Points*, Luke Nambi Mohanam and Amanda J. Holton **Journal of Chemical Education** 2020 97 (11), 4044-4048 DOI: 10.1021/acs.jchemed.0c00050 . Copyright 2020 American Chemical Society. The material in this chapter is based upon work supported by the National Science Foundation under OAC-1835909.

5.1 Chapter Motivations

Even for cutting-edge electronic structure calculation, the chemical intuition of the type and strength of intermolecular interactions is important. From carrying out electronic structure calculations and leading discussion sessions, improving student's grasp of this aspect of chemical intuition has been targeted as this concept is important for their future coursework and research career. An active learning activity was designed, implemented and evaluated for this purpose.

5.2 Introduction

A significant portion of STEM courses is devoted to the study of molecules and their interactions. [113, 114, 115] Understanding intermolecular interactions (IMFs) is important in reaction selectivity and behavior of mixtures in synthetic inorganic/organic chemistry, analytical physical chemistry, chemical engineering and food science.

Students in general chemistry are expected to understand how hydrogen bonding, dipole-dipole interactions and London Dispersion forces interact to affect the following molecular properties: boiling point, melting point viscocity, capillary action, and vapor pressure.[115] This game uses boiling points as a reasonable proxy for overall strengths of intermolecular forces. It serves as a more intuitive option than vaporization enthalpy and has fewer examples of anomalous behavior than melting temperature, viscocity and surface tension.[113, 114, 115, 116].

Additionally, students learn to rank the strength of IMFs in molecular compounds in three manners. The simplest is to rank the strength of a single IMF. For example, they should state that a molecule with a “larger electron cloud” has more LDFs and therefore a higher boiling point. A more advanced comparison requires students to compare multiple IMFs. For example, they should state that for similarly sized molecules, the species with dipole–dipole forces will have greater IMFs than the species without. The most difficult comparison asks students to compare molecules of different sizes along with different types of IMFs.

The problem of students misunderstanding IMFs or holding misconceptions about IMFs has been widely identified from the high school level and studied with interventions unsuitable for a large undergraduate general chemistry lecture. [117, 118, 119, 120, 121] The strength of LDFs, Dipole-Dipole and Hydrogen bonding have significant overlap and often classroom examples are often cherry-picked for a simplistic view of these interactions. [113, 114, 115] Students typically have difficulty in describing why a large molecule with only LDFs may have a higher boiling point than a small molecule which contains a weak dipole-dipole interaction. Inquiry activities can be a powerful way to develop student skills in this area [116, 122], but such activities can be prohibitively costly and difficult to do in large classrooms or when little resources are available.

Games have been used to teach chemistry for a century with a range of different goals for the game and the student.[123, 124] With some of these goals, some games may be complex and require significant preparation,[125] or technology.[126, 127, 128, 129] However, among these games, many have used simple card and board games to teach students chemical concepts as a “time-effective” measure to communicate nuanced chemical concepts.[130, 131, 132] These games cover a wide array of topics including bond formation, [133] organic synthesis, [134] building lewis structures [135] and building reactions.[136]

An intuition surrounding the strength and scale of the IMFs could improve students’ ability to apply the concepts previously discussed. This can be achieved through interactive exposure to a selection of boiling points of various pure substances and encouraging discussions about IMFs. [122] Here, we describe such a method of active exposure using a 3 player card game that can be played in any sized classroom with little expense.

5.3 Design Principles for Chemistry-based Games

To ensure chemistry discussion is the central mechanic of the game, the design principles are notably similar to and different from other chemistry games. [137, 138, 139, 140]

- (I) The game is based on tangible observables to increase discussion ease compared to more abstract concepts.
- (II) The game facilitates discussion through inclusion of a variety of species whose chemistry can be quickly grasped by the students.[138]
- (III) The game rules include straightforward chemistry that will lead to discussions [139] instead of memorization of external references or answer keys.[140] Requiring answer keys may shut down discussions.[138] Effective discussions require that students have been previously introduced to the topics through lecture or otherwise.
- (IV) Students have direct control over the difficulty. Having a fixed difficulty level may leave some students behind while not challenging others.
- (V) There are strategic elements in the game. There should be penalties for mistakes and rewards for risks to encourage students to flesh out their thought process. [140] It is, of course, important to avoid demoralizing students, [137, 138] but with careful thought appropriate consequences can be designed.

5.4 Game Description

5.4.1 Card Design

In this card-discard game, 36 cards (Table 5.1) containing 30 different molecules were carefully selected. Each card represents a molecule and includes the name, the chemical formula, boiling point (at atmospheric pressure, rounded off to whole numbers), and the Lewis structure drawn to show geometry. Boiling points were included to fulfill design principle (I).

As part of design principle (II), these molecules have a reasonable range of VSEPR and Lewis structures that can be easily interpreted by the students. The 30 molecules were selected specifically to avoid more complicated interactions (for example, significant zwitterion formation).

5.4.2 Game Play

The game is played in a group of three, and the game ends when one person wins by being the first to discard all cards. To start the game, each player is dealt seven cards.

The game play consists of a series of rounds. Each round starts by declaring one player "the arbiter". The remaining two players then secretly choose a card from their own hands to pass to the arbiter (face down). If no suitable comparison between the cards can be made, the arbiter may ask for a replacement card from one or both of the players. The arbiter will then place the two cards (face up) in front of the players and announce a comparison of the arbiter's choosing: The options are stronger/weaker hydrogen bonding/dipole-dipole attraction/LDFs. For example, the arbiter could ask, "Which card has stronger LDFs?" The two players then race to tap the correct card.

Table 5.1: Molecules and boiling points used in the IMF challenge

Molecule	Number of cards	Boiling Point(K)	Number of electrons	Molecular Dipole
HF	1	293	10	hydrogen bonding
HCl	1	188	18	present
HBr	1	206	36	present
HI	1	238	54	present
H_2O	3	373	10	hydrogen bonding
H_2S	1	213	18	present
H_2Se	1	232	36	present
NH_3	3	240	10	hydrogen bonding
PH_3	1	186	18	present
H_2	1	20	2	absent
N_2	1	77	14	absent
O_2	2	90	16	absent
F_2	2	85	18	absent
Cl_2	1	239	34	absent
Br_2	1	332	70	absent
I_2	1	457	106	absent
PCl_3	1	349	66	present
PCl_5	1	440	100	absent
SF_4	1	235	52	present
SF_6	1	222	70	absent
$SOCl_2$	1	348	58	present
OCH_2	1	254	16	present
CO_2	1	217	22	absent
HCN	1	299	14	present
CH_3Cl	1	249	26	present
CH_2Cl_2	1	313	42	present
$CHCl_3$	1	334	58	present
CCl_4	1	350	74	absent

Once the players have tapped the cards, the consequences of the taps must then be resolved as described below to end the round.

If only one player taps the correct card, they will discard an additional card of their choosing from their hand. The player who taps the incorrect card will draw an additional card from the deck.

If both players tap the correct card, the player who tapped the card fastest will discard an additional card from their hand. The player that taps the wrong card must draw a card from the deck.

The arbiter may inadvertently ask for an impossible comparison. The two most common mistakes are asking for a comparison of hydrogen bonding or dipole–dipole forces when they are absent in both molecules. Additionally, the three comparisons that are nearly impossible to make at a general chemistry knowledge level should be announced to the students, and it is the arbiter's responsibility to avoid these types of comparisons (see Table 5.2). If the arbiter asks for an impossible comparison, then the players can call “impossible”. The arbiter must then draw two cards. If a player taps a card for an impossible comparison, they must draw one card.

After all tap effects have been resolved, both cards used in the comparison are discarded. This ends the round.

The arbiter role rotates to the left, and a new round starts. Play continues until one player does not have any cards. In the unlikely scenario where the last two cards given to the arbiter have no reasonable comparison, the arbiter may exchange one card for a card from the deck.

A standard game takes approximately 10 min. If a longer game is preferred, more cards can be dealt during the setup. Games may also take longer if the group is especially prone to

Table 5.2: Impossible Comparisons

Cards	IMF	Reason
HBr, Cl_2	LDF	Though they have similar numbers of electrons, their LDFs are significantly different.
SF_4, HI	LDF	Sulfur tetrafluoride has significantly weaker LDFs, compared to what might be expected for a molecule with 52 electrons, due to F's high electronegativity.
SF_4, HI	Dipole-dipole	Given the molecular dipole trends, it is quite impossible to estimate this comparison.
OCH_2, HCN	Dipole-dipole	Students would not be expected to compare these dipoles without heavy guidance.
OCH_2, HCN	LDF	The polarizability of HCN is the reason for the higher boiling point even though they have similar numbers of electrons.

mistakes.

5.4.3 Game Design Rational

Design principles III and IV are fulfilled; discussing the answer is required if an answer key is not given, and the arbiter can control the difficulty of the activity. This can be cause for disagreement in the groups, and in order to ensure that more vocal students do not “win” simply through confidence and persuasion, the students were encouraged to ask for guidance from the TAs, tutors, and course instructor. Guidance was then offered in the form of guiding the conversation until the correct consensus was reached.

Design principle V is fulfilled by the race to tap and the rules for card drawing/discarding based on correctness.

While the boiling point given on the cards may assist the students in discussion, the boiling points may also help build a sense of scale about the different IMFs. In particular, some cases promote critical thinking surrounding trends beyond pattern recognition. For example, they must appropriately weight the very high LDFs present in a large molecule against a weak dipole in a smaller molecule. Similarly, they may have to appropriately judge the combination of a dipole and large LDFs against weak hydrogen bonding.

One example illustrative of critical thinking is as follows. Oxygen’s boiling point is 90.2 K while fluorine’s is 85.0 K. When asked to compare the LDFs in oxygen and fluorine, one student might tap fluorine and say, “Fluorine has more electrons and therefore has more LDFs.” The other student might tap oxygen and say, “Oxygen has the higher boiling point, and only LDFs are present; therefore, oxygen must have the higher LDFs.” The students would then need to notice and discuss the deviation from the trend and discuss why this exists.

5.4.4 Classroom use of Game

Two days after a lecture on IMFs, a class was introduced to the game, which took 10 minutes. Groups were provided one set of cards, and each student received a rule sheet and a worksheet. They were given 20 minutes to play.

The for-credit worksheet was used to promote active engagement. This has been shown to improve outcomes in classes largely comprised of non-major students with low motivation.[141] The worksheet instructs the arbiter of each turn to record the Lewis and VSEPR structure of the cards played, the IMF comparison chosen and the correct card. The design of the worksheet was not evaluated in this study to avoid reducing the n-values by introducing more variables.

5.5 Evaluation

IRB approval for studies in the professor's classes is obtained.

The pre-/post-activity test (see appendix D) consists of six 4-option multiple choice questions and the student's student ID number, allowing the use of the two-tailed paired t-test in evaluating the significance of the results. Students were given 5 minutes to respond to the test immediately before the activity(after the pre-activity lecture) and immediately after the activity.

Accurate Lewis and VSEPR structures were shown to isolate the topic to IMFs. The molecules chosen for the test are not part of the 30 molecules used in the card game. While some of these molecules have more complicated intermolecular interactions, the questions were designed to be answered at the freshman level.

The student's average score on the test improved from 3.48/6 ($\sigma^2 = 2.41$) to 3.97/6 ($\sigma^2 =$

2.35) after the activity ($p < 0.001$). Figure 5.1 shows the breakdown by question.

Question 1 had the number of correct responses increased significantly from 69.1% to 83.5% ($p < 0.001$). These students were better able to recognize the stronger LDFs in neon compared to helium after the activity.

Question 2 had the number of correct responses increased non-significantly from 60.6% to 61.9% ($p = 0.70$). After the activity, students were not significantly better or worse at recognizing that methylimine is “capable” of hydrogen bonding as a cause for its higher boiling point compared to ethene. Methylimine was picked as a “unseen” molecule since typical questions do not include unstable molecules. This suggests that the other significant results after the activity are caused by the activity itself, and also suggests the importance of presenting a range of desired functional groups in the classroom for the understanding of IMFs.

Question 3 had the number of correct responses increased non-significantly from 63.1% to 67.4% ($p = 0.22$). After the activity, students were not significantly better or worse at recognizing that the *cis-* isomer of dichloro-difluroxenon has stronger dipole-dipole interactions than the *trans-* isomer. This suggests that the game would be better if it included more molecular geometries.

In contrast, Question 4 had the number of correct responses increased significantly from 55.1% to 66.1% ($p = 0.0020$). After the activity, these students were significantly better at recognizing that the *cis-* isomer of dichloro-ethene has stronger dipole-dipole interactions than the *trans-* isomer, which would imply that even with the limited number of molecular geometries, students were able to improve in this concept. It is unexpected, given question 3 and 4 which test similar principles, that improvement would be seen on only question 4. Further development of a deeper analytical instrument would be required to elucidate this difference.

Question 5 had the number of correct responses increased significantly from 69.1% to 76.3% ($p = 0.0029$). After the activity, these students are significantly better at recognising that glycerol has more hydrogen bonding than pentanol as a cause for glycerol having the higher boiling point. This is unexpected given that no alcohols are present in the card game, and the result for question 2. The data could be confounded by the students being given the same time limit to answer all 6 questions. Being exposed to these molecular geometries through other resources might also be a factor.

Question 6 had the number of correct responses increased significantly from 30.9% to 41.5% ($p = 0.0015$). After the activity, students are significantly better at recognising that sulfur trioxide does not have dipole-dipole interactions, but has significantly stronger LDFs when compared to sulfur dioxide's weaker LDFs and dipole-dipole interaction, as a cause for sulfur trioxide having the higher boiling point. Sulfur trioxide forms dative bonds with itself, and thus is another "unseen" molecule. Question 6 is by far the question with the poorest performance, supporting the discussion in the introduction. However, the significance of improvement rivals question 1. Given the results of the other questions, this demonstrates that the game is useful for developing critical thinking skills in these students without hindering their ability to recognize simple trends.

5.6 Limitations of the study

To keep the card game focused, only molecular compounds are included. A different set of molecules would be required for focusing on ionic or metal-ligand interactions, DNA, drugs-receptor interactions, dyes, melting points, physical states, solubility, surface tension or viscosity. A document containing the cards can be found in the link in appendix D and can be edited as desired to include more or different concepts of the instructor's choosing. Due to limitations determining an appropriate control group, longitudinal knowledge and

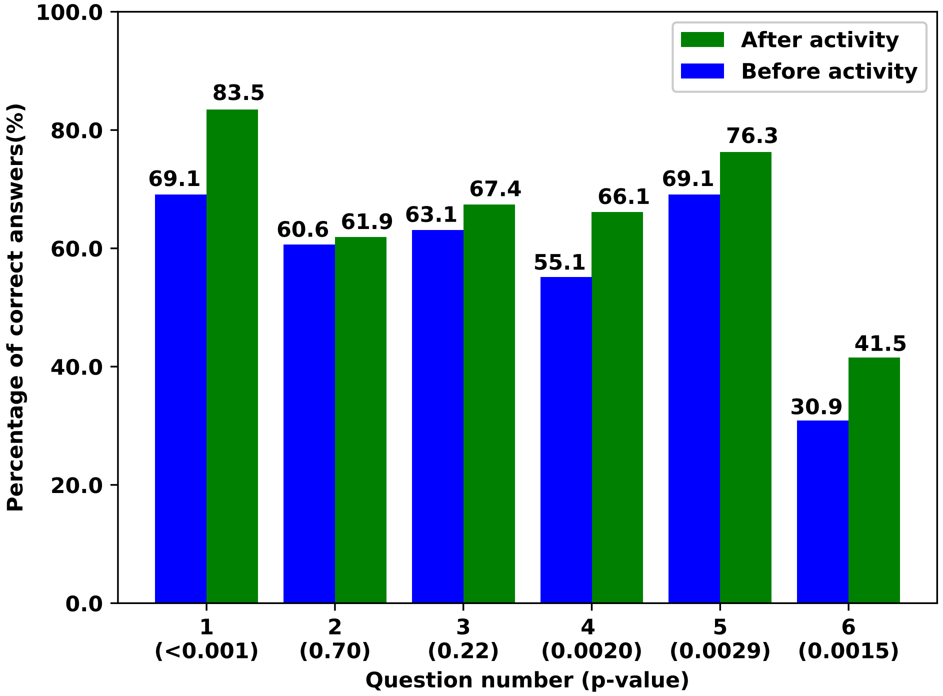

Figure 5.1: Percentage of correct answers from 236 matched responses (two-tailed paired t-test p-values in parentheses)

application retention were not tested. Other sections of the course were taught online, or with a significantly different instructor and student population. Any possible identified control group would have had significant confounding variables. Therefore the choice was made to rely only on the pre and post test paired t-tests for analysis.

5.7 Conclusions

The design objective for the IMF game discussed here has been met. Given only 20 minutes of gameplay, the students improved on several of the tested outcomes. It is important to note that this game should be integrated into a larger lesson on IMFs and is not meant to be a stand-alone introduction. Though this particular trial was on a limited set of cards, further iterations of the game could be played with other card choices aimed to improve other IMF learning outcomes not discussed in this targeted trial.

5.8 Acknowledgements

This material is based upon work supported by the National Science Foundation under OAC-1835909.

The authors would like to thank Filipp Furche for helpful discussions.

The authors would like to thank Zachary Thammavongsy for useful advice.

The following were enrolled in UCI at the time of the activity, either as graduate or undergraduate students:

The cards were sliced by hand with the help of Victoria Lim.

The following helped facilitate the activity and manage the pre-activity/post-activity test (ordered by last names):

- Teaching Assistants- Stanley C. Hiew, Samuel A. Moehring, Megan E. Screen
- Learning Assistants- Silu Chen, Linrui Li, Shayla Oliver, Natalie Samaan, Shilin Wang
- Tutors- Andrew Lee, Andrew Tang, Avin Seneth Wijayaweera

Bibliography

[1] F. Furche: On the density matrix based approach to time-dependent density functional response theory, *J. Chem. Phys.* **114**(14), 5982–5992, 2001.

[2] S. M. Parker, D. Rappoport, and F. Furche: Quadratic Response Properties from TDDFT: Trials and Tribulations, *J. Chem. Theo. Comput.* **14**(2), 807–819, 2018.

[3] F. Furche, B. T. Krull, B. D. Nguyen, and J. Kwon: Accelerating molecular property calculations with nonorthonormal Krylov space methods, *J. Chem. Phys.* **144**(17), 174105, 2016.

[4] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński, and J. M. Yu: TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations, *J. Chem. Phys.* **152**(18), 184107, 2020.

[5] L. N. Mohanam, S. Bekoe, Z. Shen, N. George, and F. Furche: *libkrylov*, a Modular Open-Source Software Library for Extremely Large Eigenvalue and Linear Problems, 2020, https://doi.org/10.5281/zenodo.3974054.

[6] E. Anderson, Z. Bai, C. Bischof, L. S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen: LAPACK Users' Guide, Society for Industrial and Applied Mathematics, third edition, 1999.

[7] M. Hochstenbach and Y. Notay: The Jacobi–Davidson method, *GAMM-Mitteilungen* **29**(2), 368–382, 2006.

[8] G. L. G. Sleijpen and H. A. Van der Vorst: A Jacobi–Davidson Iteration Method for Linear Eigenvalue Problems, *SIAM Review* **42**(2), 267–293, 2000.

[9] M. Genseberger and G. L. G. Sleijpen: Alternative correction equations in the Jacobi-Davidson method, *Num. Lin. Alg. Appl.* **6**(3), 235–253, 1999.

[10] R. Bauernschmitt and R. Ahlrichs: Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory, *Chem. Phys. Lett.* **256**(4), 454–464, 1996.

[11] C. Cohen-Tannoudji and B. Laloë: Quantum Mechanics Volume 1, Hermann, 2005.

[12] T. Fransson, I. E. Brumboiu, M. L. Vidal, P. Norman, S. Coriani, and A. Dreuw: XABOOM: An X-ray Absorption Benchmark of Organic Molecules Based on Carbon, Nitrogen, and Oxygen 1s → π^* Transitions, *J. Chem. Theo. Comput.* **17**(3), 1618–1637, 2021.

[13] E. Tapavicza, G. D. Bellchambers, J. C. Vincent, and F. Furche: Ab initio non-adiabatic molecular dynamics, *Phys. Chem. Chem. Phys.* **15**, 18336–18348, 2013.

[14] S. Grimme: Density functional calculations with configuration interaction for the excited states of molecules, *Chemical Physics Letters* **259**(1), 128–137, 1996.

[15] Y. Sadd: Iterative Methods for Sparse Linear Systems, Society for Industrial and Applied Mathematics, second edition, 2003.

[16] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs: Auxiliary basis sets to approximate Coulomb potentials, *Chem. Phys. Lett.* **240**(4), 283–290, 1995.

[17] A. Szabo and N. S. Ostlund: Modern Quantum Chemistry, Introduction to Advanced Electronic Structure Theory, Dover, second edition, 1996.

[18] G. Golub and C. Van Loan: Matrix Computations, Johns Hopkins Studies in Atlantic History & Culture, Johns Hopkins University Press, 1983.

[19] E. R. Davidson: The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, *J. Comput. Phys.* **17**(1), 87–94, 1975.

[20] C. Huang, W. Liu, Y. Xiao, and M. R. Hoffmann: iVI: An iterative vector interaction method for large eigenvalue problems, *J. Comput. Chem.* **38**(29), 2481–2499, 2017.

[21] M. Hubert, H. J. A. Jensen, and E. D. Hedegărd: Excitation Spectra of Nucleobases with Multiconfigurational Density Functional Theory, *J. Phys. Chem. A* **120**(1), 36–43, 2016.

[22] J. P. Perdew, K. Burke, and M. Ernzerhof: Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865–3868, 1996.

[23] F. Weigend and R. Ahlrichs: Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *Phys. Chem. Chem. Phys.* **7**, 3297, 2005.

[24] J. P. Perdew, M. Ernzerhof, and K. Burke: Rationale for mixing exact exchange with density functional approximations, *J. Chem. Phys.* **105**(22), 9982–9985, 1996.

[25] C. Adamo and V. Barone: Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.* **110**(13), 6158–6170, 1999.

[26] J. C. Tully: Molecular dynamics with electronic transitions, *J. Chem. Phys.* **93**(2), 1061–1071, 1990.

[27] S. M. Parker, S. Roy, and F. Furche: Multistate hybrid time-dependent density functional theory with surface hopping accurately captures ultrafast thymine photodeactivation, *Phys. Chem. Chem. Phys.* **21**, 18999–19010, 2019.

[28] E. Tapavicza, A. M. Meyer, and F. Furche: Unravelling the details of vitamin D photosynthesis by non-adiabatic molecular dynamics simulations, *Phys. Chem. Chem. Phys.* **13**, 20986–20998, 2011.

[29] J. C. Vincent, M. Muuronen, K. C. Pearce, L. N. Mohanam, E. Tapavicza, and F. Furche: That Little Extra Kick: Nonadiabatic Effects in Acetaldehyde Photodissociation, *J. Phys. Chem. Lett.* **7**(20), 4185–4190, 2016.

[30] M. Muuronen, S. M. Parker, E. Berardo, A. Le, M. A. Zwijnenburg, and F. Furche: Mechanism of photocatalytic water oxidation on small TiO2 nanoparticles, *Chem. Sci.* **8**, 2179–2183, 2017.

[31] S. Roy, S. Ardo, and F. Furche: 5-Methoxyquinoline Photobasicity Is Mediated by Water Oxidation, *J. Phys. Chem. A* **123**(31), 6645–6651, 2019.

[32] A. Brown, C. M. Kemp, and S. F. Mason: Electronic absorption, polarised excitation, and circular dichroism spectra of [5]-helicene (dibenzo[c,g]phenanthrene), *J. Chem. Soc. A* pp. 751–755, 1971.

[33] J. C. del Valle and J. Catalán: Kasha's rule: a reappraisal, *Phys. Chem. Chem. Phys.* **21**, 10061–10069, 2019.

[34] T. Itoh: Fluorescence and Phosphorescence from Higher Excited States of Organic Molecules, *Chem. Rev.* **112**(8), 4541–4568, 2012.

[35] K. Veys and D. Escudero: Computational Protocol To Predict Anti-Kasha Emissions: The Case of Azulene Derivatives, *J. Phys. Chem. A* **124**(36), 7228–7237, 2020.

[36] A. Prlj, T. Begušić, Z. T. Zhang, G. C. Fish, M. Wehrle, T. Zimmermann, S. Choi, J. Roulet, J.-E. Moser, and J. Vaníček: Semiclassical Approach to Photophysics Beyond Kasha's Rule and Vibronic Spectroscopy Beyond the Condon Approximation. The Case of Azulene, *J. Chem. Theo. Comput.* **16**(4), 2617–2626, 2020.

[37] Y. Amatatsu and Y. Komura: Reaction coordinate analysis of the $S_1 - S_0$ internal conversion of azulene, *J. Chem. Phys.* **125**(17), 174311, 2006.

[38] S. Murata, C. Iwanaga, T. Toda, and H. Kokubun: Fluorescence and radiationless transitions from the second excited states of azulene derivatives, *Berichte der Bunsengesellschaft für physikalische Chemie* **76**(11), 1176–1183, 1972.

[39] H. J. Griesser and U. P. Wild: The energy gap dependence of the radiationless transition rates in azulene and its derivatives, *Chem. Phys.* **52**(1), 117–131, 1980.

[40] D. Huppert, J. Jortner, and P. M. Rentzepis: Azulene Revisited: Picosecond Decay of the S_1 State in the Gas Phase and in Solution, *Isr. J. Chem.* **16**(4), 277–282, 1977.

[41] D. Huppert, J. Jortner, and P. M. Rentzepis: Laser Excited Emission Spectroscopy of Azulene in the Gas Phase, *J. Chem. Phys.* **56**(10), 4826–4833, 1972.

[42] E. W.-G. Diau, S. De Feyter, and A. H. Zewail: Direct observation of the femtosecond nonradiative dynamics of azulene in a molecular beam: The anomalous behavior in the isolated molecule, *J. Chem. Phys.* **110**(20), 9785–9788, 1999.

[43] M. Orenstein, S. Kimel, and S. Speiser: Laser excited $\mathrm{S}_2 \rightarrow \mathrm{S}_1$ and $\mathrm{S}_1 \rightarrow \mathrm{S}_0$ emission spectra and the $\mathrm{S}_2 \rightarrow \mathrm{S}_n$ absorption spectrum of azulene in solution, *Chem. Phys. Lett.* **58**(4), 582–585, 1978.

[44] J. M. Friedman and R. M. Hochstrasser: Moderately high resolution fluorescence spectrum of the $\mathrm{S}_1 \rightarrow \mathrm{S}_0$ transition of azulene, *Chem. Phys.* **6**(2), 145–154, 1974.

[45] G. Viswanath and M. Kasha: Confirmation of the Anomalous Fluorescence of Azulene, *J. Chem. Phys.* **24**(3), 574–577, 1956.

[46] G. Binsch, E. Heilbronner, R. Jankow, and D. Schmidt: On the fluorescence anomaly of azulene, *Chem. Phys. Lett.* **1**(4), 135–138, 1967.

[47] B.-C. Wang, Y.-S. Lin, J.-C. Chang, and P.-Y. Wang: Theoretical studies of azulene and its derivatives, *Can. J. Chem.* **78**(2), 224–232, 2000.

[48] G. R. Hunt and I. G. Ross: Spectrum of azulene: Part II. The 7000-A and 3500-A absorption systems, *J. Mol. Spectrosc.* **9**, 50–78, 1962.

[49] G. Eber, F. Grüneis, S. Schneider, and F. Dörr: Dual fluorescence emission of azulene derivatives in solution, *Chem. Phys. Lett.* **29**(3), 397–404, 1974.

[50] M. Kasha: Characterization of electronic transitions in complex molecules, *Discuss. Faraday Soc.* **9**, 14–19, 1950.

[51] F. Weigend: Accurate Coulomb-fitting basis sets for H to Rn, *Phys. Chem. Chem. Phys.* **8**, 1057–1065, 2006.

[52] M. Allen, D. Tildesley, and T. Allen: Computer Simulation of Liquids, Oxford Science Publ, Clarendon Press, 1989.

[53] Q. Ou, G. D. Bellchambers, F. Furche, and J. E. Subotnik: First-order derivative couplings between excited states from adiabatic TDDFT response theory, *J. Chem. Phys.* **142**(6), 064114, 2015.

[54] M. Barbatti: Velocity Adjustment in Surface Hopping: Ethylene as a Case Study of the Maximum Error Caused by Direction Choice, *J. Chem. Theo. Comput.* **17**(5), 3010–3018, 2021.

[55] S. Blanes, F. Casas, and J. Ros: Improved High Order Integrators Based on the Magnus Expansion, *BIT Numer. Math.* **40**, 434–450, 2000.

[56] A. Jain, E. Alguire, and J. E. Subotnik: An Efficient, Augmented Surface Hopping Algorithm That Includes Decoherence for Use in Large-Scale Simulations, *J. Chem. Theo. Comput.* **12**(11), 5256–5268, 2016.

[57] L. Wang, D. Trivedi, and O. V. Prezhdo: Global Flux Surface Hopping Approach for Mixed Quantum-Classical Dynamics, *J. Chem. Theo. Comput.* **10**(9), 3598–3605, 2014.

[58] W. J. Evans: Tutorial on the Role of Cyclopentadienyl Ligands in the Discovery of Molecular Complexes of the Rare-Earth and Actinide Metals in New Oxidation States, *Organometallics* **35**, 3088, 2016.

[59] D. H. Woen and W. J. Evans: Chapter 293 - Expanding the +2 Oxidation State of the Rare-Earth Metals, Uranium, and Thorium in Molecular Complexes, in J.-C. G. Bünzli and V. K. Pecharsky, editors, Including Actinides, volume 50 of *Handbook on the Physics and Chemistry of Rare Earths*, pp. 337–394, Elsevier, 2016.

[60] G. Meyer: Reduced Halides of the Rare-Earth Elements, *Chem. Rev.* **88**, 93, 1988.

[61] M. N. Bochkarev: Molecular compounds of "new" divalent lanthanides, *Coord. Chem. Rev.* **248**, 835, 2004.

[62] G. Meyer: The Divalent State in Solid Rare-Earth Metal Halides, in The Rare Earth Elements, John Wiley & Sons, 2012.

[63] F. Nief: Chapter 246 Molecular Chemistry of the Rare-Earth Elements in Uncommon Low-Valent States, *Handbook on the Physics and Chemistry of Rare Earths* **40**, 241–300, 2010.

[64] P. B. Hitchcock, M. F. Lappert, L. Maron, and A. V. Protchenko: Lanthanum does form stable molecular compounds in the +2 oxidation state, *Angew. Chem., Int. Ed.* **47**, 1488, 2008.

[65] C. T. Palumbo, L. E. Darago, C. J. Windorff, J. W. Ziller, and W. J. Evans: Trimethylsilyl versus Bis(trimethylsilyl) Substitution in Tris(cyclopentadienyl) Complexes of La, Ce, and Pr: Comparison of Structure, Magnetic Properties, and Reactivity, *Organometallics* **37**, 900, 2018.

[66] M. R. MacDonald, J. E. Bates, J. W. Ziller, F. Furche, and W. J. Evans: Completing the series of +2 ions for the lanthanide elements: synthesis of molecular complexes of Pr^2+, Gd^2+, Tb^2+, and Lu^2+, *J. Am. Chem. Soc.* **135**, 9857, 2013.

[67] M. E. Fieser, M. R. MacDonald, B. T. Krull, J. E. Bates, J. W. Ziller, F. Furche, and W. J. Evans: Structural, spectroscopic, and theoretical comparison of traditional vs recently discovered Ln^{2+} ions in the $[K(2.2.2\text{-cryptand})][(C_5H_4SiMe_3)_3Ln]$ complexes: the variable nature of Dy^{2+} and Nd^{2+}, *J. Am. Chem. Soc.* **137**, 369, 2015.

[68] M. R. MacDonald, J. W. Ziller, and W. J. Evans: Synthesis of a crystalline molecular complex of Y^{2+}, $[(18\text{-crown-6})K][(C_5H_4SiMe_3)_3Y]$, *J. Am. Chem. Soc.* **133**, 15914, 2011.

[69] M. R. MacDonald, J. E. Bates, M. E. Fieser, J. W. Ziller, F. Furche, and W. J. Evans: Expanding Rare-Earth Oxidation State Chemistry to Molecular Complexes of Holmium(II) and Erbium(II), *J. Am. Chem. Soc.* **134**, 8420, 2012.

[70] R. R. Langeslay, M. E. Fieser, J. W. Ziller, F. Furche, and W. J. Evans: Synthesis, structure, and reactivity of crystalline molecular complexes of the $\{[C_5H_3(SiMe_3)_2]_3Th\}^{1-}$ anion containing thorium in the formal +2 oxidation state, *Chem. Sci.* **6**, 517, 2015.

[71] M. R. MacDonald, M. E. Fieser, J. E. Bates, J. W. Ziller, F. Furche, and W. J. Evans: Identification of the +2 Oxidation State for Uranium in a Crystalline Molecular Complex, $[K(2.2.2\text{-Cryptand})][(C_5H_4SiMe_3)_3U]$, *J. Am. Chem. Soc.* **135**, 13310, 2013.

[72] C. J. Windorff, M. R. MacDonald, K. R. Meihaus, J. W. Ziller, J. R. Long, and W. J. Evans: Expanding the Chemistry of Molecular U^{2+} Complexes: Synthesis, Characterization, and Reactivity of the $\{[C_5H_3(SiMe_3)_2]_3U\}^-$ Anion, *Chem. Eur. J.* **22**, 772, 2016.

[73] C. J. Windorff, G. P. Chen, J. N. Cross, W. J. Evans, F. Furche, A. J. Gaunt, M. T. Janicke, S. A. Kozimor, and B. L. Scott: Identification of the Formal +2 Oxidation State of Plutonium: Synthesis and Characterization of $\{Pu(II)[C_5H_3(SiMe_3)_2]_3\}^-$, *J. Am. Chem. Soc.* **139**, 3970, 2017.

[74] J. Su, C. J. Windorff, E. R. Batista, W. J. Evans, A. J. Gaunt, M. T. Janicke, S. A. Kozimor, B. L. Scott, D. H. Woen, and P. Yang: Identification of the Formal +2 Oxidation State of Neptunium: Synthesis and Structural Characterization of $\{Np^{II}[C_5H_3(SiMe_3)_2]_3\}^{1-}$, *J. Am. Chem. Soc.* **140**, 7425, 2018.

[75] Y. K. Gun'ko, P. B. Hitchcock, and M. F. Lappert: Activation of a C-O bond by reaction of a tris(cyclopentadienyl)lanthanide complex with an alkali metal in dimethoxyethane (DME); crystal structures of $[Nd\{\eta\text{-}C_5H_3(SiMe_3)_2\text{-}1,3\}_2(\mu\text{-}OMe)_2Li(DME)]$ and $[Ce(\eta\text{-}C_5H_3t\text{Bu}_2\text{-}1,3)_2(\mu\text{-}OMe)_2]$, *J. Organomet. Chem.* **499**, 213, 1995.

[76] M. C. Cassani, D. J. Duncalf, and M. F. Lappert: The First Example of a Crystalline Subvalent Organolanthanum Complex: $[K([18]\text{crown-6})\text{-}(\eta^2\text{-}C_6H_6)_2][(LaCp^{tt}_2)_2(\mu\text{-}\eta^6{:}\eta^6\text{-}C_6H_6)] \cdot 2C_6H_6$ ($Cp^{tt}= \eta^5\text{-}C_5H_3Bu^t_2\text{-}1,3$), *J. Am. Chem. Soc.* **120**, 12958, 1998.

[77] F. Jaroschik, F. Nief, X. F. Le Goff, and L. Ricard: Isolation of stable organodysprosium(II) complexes by chemical reduction of dysprosium(III) precursors, *Organometallics* **26**, 1123, 2007.

[78] F. Jaroschik, A. Momin, F. Nief, X.-F. Le Goff, G. B. Deacon, and P. C. Junk: Dinitrogen Reduction and C-H Activation by the Divalent Organoneodymium Complex $[(C_5H_2\,{}^tBu_3)_2Nd(\mu\text{–}I)K([18]crown\text{-}6)]$, *Angew. Chem.* **121**, 1137, 2009.

[79] H. S. La Pierre, A. Scheurer, F. W. Heinemann, W. Hieringer, and K. Meyer: Synthesis and characterization of a uranium(II) monoarene complex supported by delta backbonding, *Angew. Chem., Int. Ed.* **53**, 7158, 2014.

[80] M. E. Fieser, M. G. Ferrier, J. Su, E. Batista, S. K. Cary, J. W. Engle, W. J. Evans, J. S. Lezama Pacheco, S. A. Kozimor, A. C. Olson, A. J. Ryan, B. W. Stein, G. L. Wagner, D. H. Woen, T. Vitova, and P. Yang: Evaluating the electronic structure of formal Ln^{II} ions in $Ln^{II}(C_5H_4SiMe_3)_3^{1-}$ using XANES spectroscopy and DFT calculations, *Chemical Science* **8**, 6076, 2017.

[81] K. R. Meihaus, M. E. Fieser, J. F. Corbey, W. J. Evans, and J. R. Long: Record High Single-Ion Magnetic Moments Through $4f^n5d^1$ Electron Configurations in the Divalent Lanthanide Complexes $[(C_5H_4SiMe_3)_3Ln]^-$, *J. Am. Chem. Soc.* **137**, 9855, 2015.

[82] J. F. Corbey, D. H. Woen, C. T. Palumbo, M. E. Fieser, J. W. Ziller, F. Furche, and W. J. Evans: Ligand Effects in the Synthesis of Ln^{2+} Complexes by Reduction of Tris(cyclopentadienyl) Precursors Including C-H Bond Activation of an Indenyl Anion, *Organometallics* **34**, 3909, 2015.

[83] D. N. Huh, C. M. Kotyk, M. Gembicky, A. L. Rheingold, J. W. Ziller, and W. J. Evans: Synthesis of rare-earth-metal-in-cryptand dications, $[Ln(2.2.2\text{-}cryptand)]^{2+}$, from Sm^{2+}, Eu^{2+}, and Yb^{2+} silyl metallocenes $(C_5H_4SiMe_3)_2Ln(THF)_2$, *Chem. Commun.* **53**, 8664, 2017.

[84] J. W. Lauher and R. Hoffmann: Structure and chemistry of bis(cyclopentadienyl)-MLn complexes, *J. Am. Chem. Soc.* **98**, 1729, 1976.

[85] B. E. Bursten, L. F. Rhodes, and R. J. Strittmatter: Bonding in Tris(η^5-cyclopentadienyl) Actinide Complexes. 2. On the Ground Electronic Configurations of "Base-Free" Cp,An Complexes (An = Th, Pa, U, Np, Pu), *J. Am. Chem. Soc.* **111**, 2756, 1989.

[86] B. E. Bursten, L. F. Rhodes, and R. J. Strittmatter: Bonding in Tris(η^5-cyclopentadienyl) Actinide Complexes. 3. Interaction of 7-Neutral, 7-Acidic, and 7-Basic Ligands with $(T_5\text{-}C_5H_5)_3U$, *J. Am. Chem. Soc.* **111**, 2758, 1989.

[87] R. J. Strittmatter and B. E. Bursten: Bonding in Tris(η^5-Cyclopentadienyl) Actinide Complexes 0.5. A Comparison of the Bonding in Np, Pu, and Transplutonium Compounds with That in Lanthanide Compounds and a Transition-Metal Analog, *J. Am. Chem. Soc.* **113**, 552, 1991.

[88] W. W. Lukens and R. A. Andersen: Synthesis, Structure, and Reactions of (η^5-$C_5H_5)_3Zr$, *Organometallics* **14**, 3435, 1995.

[89] R. G. Denning, J. Harmer, J. C. Green, and M. Irwin: Covalency in the 4f Shell of tris-Cyclopentadienyl Ytterbium $(YbCp_3)^-$ A Spectroscopic Evaluation, *J. Am. Chem. Soc.* **133**, 20644, 2011.

[90] H. P. Beck: Notizen: NdI_2-II, eine metallisch leitende Hochdruckmodifikation NdI_2, a Metallic High Pressure Modification, *Z. Naturforsch., B: J. Chem. Sci.* **31**, 1548, 1976.

[91] N. Gerlitzki, G. Meyer, A.-V. Mudring, and J. D. Corbett: Praseodymium diiodide, PrI_2, revisited by synthesis, structure determination and theory, *J. Alloys Compd.* **380**, 211, 2004.

[92] D. H. Woen, G. P. Chen, J. W. Ziller, T. J. Boyle, F. Furche, and W. J. Evans: Solution Synthesis, Structure, and CO_2 Reduction Reactivity of a Scandium(II) Complex, $\{Sc[N(SiMe_3)_2]_3\}^{1-}$, *Angew. Chem., Int. Ed.* **56**, 2050, 2017.

[93] A. J. Ryan, L. E. Darago, S. G. Balasubramani, G. P. Chen, J. Ziller, F. Furche, and W. J. Evans: Synthesis, Structure, and Magnetism of Tris(amide) $[Ln\{N(SiMe_3)_2\}_3]^{1-}$ Complexes of the Non-traditional +2 Lanthanide Ions, *Chem. - Eur. J.* **24**, 7702, 2018.

[94] M. F. Lappert, A. Singh, J. L. Atwood, and W. E. Hunter: Use of the bis(trimethylsilyl)cyclopentadienyl ligand for stabilising early ($f^0 - f^3$) lanthanocene chlorides; X-ray structure of $[(Pr\{\eta\text{-}[C_5H_3(SiMe_3)_2]\}_2Cl)_2]$ and of isoleptic scandium and ytterbium complexes, *J. Chem. Soc., Chem. Commun.* p. 1190, 1981.

[95] H. Schumann, M. Glanz, H. Hemling, and F. Ekkehard Hahn: Metallorganische Verbindungen der Lanthanoide. 93 [1]. Tetramethylcyclopentadienyl-Komplexe ausgewhlter 4f-Elemente, *Z. Anorg. Allg. Chem.* **621**, 341, 1995.

[96] Z. W. Xie, K. L. Chui, Z. X. Liu, F. Xue, Z. Y. Zhang, T. C. W. Mak, and J. Sun: Systematic studies on the reactions of lanthanide trichlorides with Na[1,3-bis(trimethylsilyl)cyclopentadienyl]. Crystal structures of $[1,3\text{-}(Me_3Si)_2C_5H_3]_3Ln$ (Ln = La, Nd, Gd, Dy), *J. Organomet. Chem.* **549**, 239, 1997.

[97] J. K. Peterson, M. R. MacDonald, J. W. Ziller, and W. J. Evans: Synthetic Aspects of $(C_5H_4SiMe_3)_3Ln$ Rare-Earth Chemistry: Formation of $(C_5H_4SiMe_3)_3Lu$ via $[(C_5H_4SiMe_3)_2Ln]^+$ Metallocene Precursors, *Organometallics* **32**, 2625, 2013.

[98] W. J. Evans, D. B. Rego, and J. W. Ziller: Synthesis, structure, and ^{15}N NMR studies of paramagnetic lanthanide complexes obtained by reduction of dinitrogen, *Inorg. Chem.* **45**, 10790, 2006.

[99] S. D. Stults, R. A. Andersen, and A. Zalkin: Structural Studies on Cyclopentadienyl Compounds of Trivalent Cerium - Tetrameric $(Me_5H_4)_3Ce$ and Monomeric $(Me_3SiC_5H_4)_3Ce$ and $[(Me_3Si)_2C_5H_3]_3Ce$ and Their Coordination Chemistry, *Organometallics* **9**, 115, 1990.

[100] S. G. Minasian, J. L. Krinsky, J. D. Rinehart, R. Copping, T. Tyliszczak, M. Janousch, D. K. Shuh, and J. Arnold: A comparison of 4*f* vs 5*f* metal-metal bonds in $(CpSiMe_3)_3$M-ECp* (M = Nd, U; E = Al, Ga; Cp* = C_5Me_5): synthesis, thermodynamics, magnetism, and electronic structure, *J. Am. Chem. Soc.* **131**, 13767, 2009.

[101] C. J. Windorff, M. T. Dumas, J. W. Ziller, A. J. Gaunt, S. A. Kozimor, and W. J. Evans: Small-Scale Metal-Based Syntheses of Lanthanide Iodide, Amide, and Cyclopentadienyl Complexes as Analogues for Transuranic Reactions, *Inorg. Chem.* **56**, 11981, 2017.

[102] M. Dolg, H. Stoll, A. Savin, and H. Preuss: Energy-adjusted pseudopotentials for the rare earth elements, *Theor. Chim. Acta* **75**, 173, 1989.

[103] R. Gulde, P. Pollak, and F. Weigend: Error-Balanced Segmented Contracted Basis Sets of Double-ζ to Quadruple-ζ Valence Quality for the Lanthanides, *J. Chem. Theo. Comput.* **8**, 4062, 2012.

[104] A. Klamt and G. Schüürmann: COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, *J. Chem. Soc., Perkin Trans. 2* pp. 799–805, 1993.

[105] D. R. Lide: CRC Handbook of Chemistry and Physics, 81st Edition: A Ready- Reference Book of Chemical and Physical Data, CRC Press, 2008.

[106] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew: Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes, *J. Chem. Phys.* **119**(23), 12129–12137, 2003.

[107] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg: A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.* **132**(15), 154104, 2010.

[108] M. Hülsen, A. Weigand, and M. Dolg: Quasirelativistic energy-consistent 4f-in-core pseudopotentials for tetravalent lanthanide elements, *Theor. Chem. Acc.* **122**, 23–29, 2009.

[109] M. Dolg, H. Stoll, and H. Preuss: A combination of quasirelativistic pseudopotential and ligand field calculations for lanthanoid compounds., *Theor. Chim. Acta* **85**, 441–450, 1993.

[110] M. Dolg, H. Stoll, and H. Preuss: Energy-adjusted ab initio pseudopotentials for the rare earth elements, *J. Chem. Phys.* **90**(3), 1730–1734, 1989.

[111] J. E. Bates and F. Furche: Harnessing the meta-generalized gradient approximation for time-dependent density functional theory, *J. Chem. Phys.* **137**(16), 164105, 2012.

[112] D. Rappoport and F. Furche: Property-optimized Gaussian basis sets for molecular response calculations, *J. Chem. Phys.* **133**(13), 134105, 2010.

[113] G. D. Peckman and I. J. McNaught: Teaching Intermolecular Forces to First-Year Undergraduate Students, *J. Chem. Educ.* **89**(7), 955–957, 2012.

[114] M. M. Cooper, L. C. Williams, and S. M. Underwood: Student Understanding of Intermolecular Forces: A Multimodal Study, *J. Chem. Educ.* **92**(8), 1288–1298, 2015.

[115] N. J. Tro: Liquids, Solids, and Intermolecular Forces, chapter 11, p. 466, Pearson, 2018.

[116] S. Glazier, N. Marano, and L. Eisen: A Closer Look at Trends in Boiling Points of Hydrides: Using an Inquiry-Based Approach To Teach Intermolecular Forces of Attraction, *J. Chem. Educ.* **87**(12), 1336–1341, 2010.

[117] R. F. Peterson and D. F. Treagust: Grade-12 students' misconceptions of covalent bonding and structure, *J. Chem. Educ.* **66**(6), 459, 1989.

[118] M. B. Nakhleh: Why some students don't learn chemistry: Chemical misconceptions, *J. Chem. Educ.* **69**(3), 191, 1992.

[119] P. G. Jasien: Helping Students Assess the Relative Importance of Different Intermolecular Interactions, *J. Chem. Educ.* **85**(9), 1222, 2008.

[120] C. J. Luxford and S. L. Bretz: Development of the Bonding Representations Inventory To Identify Student Misconceptions about Covalent and Ionic Bonding Representations, *J. Chem. Educ.* **91**(3), 312–320, 2014.

[121] V. Kind: Beyond Appearances: Students misconceptions about basic chemical ideas, 2016.

[122] M. Ogden: An Inquiry Experience with High School Students To Develop an Understanding of Intermolecular Forces by Relating Boiling Point Trends and Molecular Structure, *J. Chem. Educ.* **94**(7), 897–902, 2017.

[123] J. V. Russell: Using Games To Teach Chemistry: An Annotated Bibliography, *J. Chem. Educ.* **76**(4), 481, 1999.

[124] M. J. Samide and A. M. Wilson: Games, Games, Games; Playing to Engage with Chemistry Concepts, *Chem. Educ.* **19**, 167–170, 2014.

[125] M. L. Clapson, B. Gilbert, V. J. Mozol, S. Schechtel, J. Tran, and S. White: ChemEscape: Educational Battle Box Puzzle Activities for Engaging Outreach and Active Learning in General Chemistry, *J. Chem. Educ.* **97**(1), 125–131, 2020.

[126] J. Winter, M. Wentzel, and S. Ahluwalia: Chairs!: A Mobile Game for Organic Chemistry Students To Learn the Ring Flip of Cyclohexane, *J. Chem. Educ.* **93**(9), 1657–1659, 2016.

[127] L. Dsilva, S. Mittal, B. Koepnick, J. Flatten, S. Cooper, and S. Horowitz: Creating custom Foldit puzzles for teaching biochemistry, *Biochem. Mol. Biol. Educ.* **47**(2), 133–139, 2019.

[128] M. F. Silva, P. M. Martins, D. C. Mariano, L. H. Santos, I. Pastorini, N. Pantuza, C. N. Nobre, and R. C. de Melo-Minardi: Proteingo: Motivation, user experience, and learning of molecular interactions in biological complexes, *Ent. Com.* **29**, 31 – 42, 2019.

[129] N. Srisawasdi and P. Panjaburee: Implementation of Game-transformed Inquiry-based Learning to Promote the Understanding of and Motivation to Learn Chemistry, *J. Sci. Educ. Technol.* **28**, 152–164, 2019.

[130] Erlina, C. Cane, and D. P. Williams: Prediction! The VSEPR Game: Using Cards and Molecular Model Building To Actively Enhance Students' Understanding of Molecular Geometry, *J. Chem. Educ.* **95**(6), 991–995, 2018.

[131] J. L. Miller, M. T. Wentzel, J. H. Clark, and G. A. Hurst: Green Machine: A Card Game Introducing Students to Systems Thinking in Green Chemistry by Strategizing the Creation of a Recycling Plant, *J. Chem. Educ.* **96**(12), 3006–3013, 2019.

[132] V. Martí-Centelles and J. Rubio-Magnieto: ChemMend: A Card Game To Introduce and Explore the Periodic Table while Engaging Students' Interest, *J. Chem. Educ.* **91**(6), 868–871, 2014.

[133] A. G. Samuelson: Card Games and Chemistry Teaching Organometallic Reactions Through Card Games, *Resonance* **23**, 915–923, 2018.

[134] S. C. Farmer and M. K. Schuman: A Simple Card Game To Teach Synthesis in Organic Chemistry Courses, *J. Chem. Educ.* **93**(4), 695–698, 2016.

[135] P. T. Bell, A. D. Adkins, R. J. Gamble, and L. D. Schultz: Enthalpy Costs of Making and Breaking Bonds: A Game of Generating Molecules with Proper Lewis Structures, *J. Chem. Educ.* **86**(4), 450, 2009.

[136] K. Gogal, W. Heuett, and D. Jaber: CHEMCompete: An Organic Chemistry Card Game To Differentiate between Substitution and Elimination Reactions of Alkyl Halides, *J. Chem. Educ.* **94**(9), 1276–1279, 2017.

[137] M. Antunes, M. A. R. Pacheco, and M. Giovanela: Design and Implementation of an Educational Game for Teaching Chemistry in Higher Education, *J. Chem. Educ.* **89**(4), 517–521, 2012.

[138] J. N. da Silva Júnior, D. E. de Andrade Uchoa, M. A. S. Lima, and A. J. Monteiro: Stereochemistry Game: Creating and Playing a Fun Board Game To Engage Students in Reviewing Stereochemistry Concepts, *J. Chem. Educ.* **96**(8), 1680–1685, 2019.

[139] R. D. Huelsmann, A. F. Vailati, L. R. de Laia, P. S. Tessaro, and F. R. Xavier: Tap It Fast! Playing a Molecular Symmetry Game for Practice and Formative Assessment of Students' Understanding of Symmetry Concepts, *J. Chem. Educ.* **95**(7), 1151–1155, 2018.

[140] C. A. Knudtson: ChemKarta: A Card Game for Teaching Functional Groups in Undergraduate Organic Chemistry, *J. Chem. Educ.* **92**(9), 1514–1517, 2015.

[141] W. H. He, A. J. Holton, and G. Farkas: Impact of partially flipped instruction on immediate and subsequent course performance in a large undergraduate chemistry course, *Comput. Educ.* **125**, 120–131, 2018.

Appendix A

Krylov Subspace Algorithm

A.1 General Notation and Lagrangian

The following notation allows for discussion of eigenvalue, linear and shifted linear problems together- Eigenvalue problems can be written as:

$$\mathbf{AX} - \mathbf{X\Omega} = \acute{\mathbf{A}}\mathbf{X} = \hat{\mathbf{A}}\mathbf{X} = \mathbf{0}$$

Linear problems can be written as:

$$\mathbf{AX} - \mathbf{P} = \acute{\mathbf{A}}\mathbf{X} - \mathbf{P} = \hat{\mathbf{A}}\mathbf{X} = \mathbf{0}$$

Shifted linear problems can be written as:

$$\mathbf{AX} - \mathbf{X}\boldsymbol{\omega} - \mathbf{P} = \acute{\mathbf{A}}\mathbf{X} - \mathbf{P} = \hat{\mathbf{A}}\mathbf{X} = \mathbf{0}$$

where the hat accent is used to summarize the entire problem to be solved, and the acute accent is used for the LHS of the problem, and the solution matrix is $\mathbf{X}$. The acute operator is linear (and sometimes described as a super operator), while the hat operation is linear only for the eigenvalue problem. (Adding a vector to another vector is not a linear operation, and is usually treated as some form of translation.) In all cases, with the finite basis for the problem, the desired solution is projected by the $\hat{\mathbf{A}}$ operator onto the zero vector in the same space.

Writing the Lagrangian of the full problems: The eigenvalue problem Lagrangian can be written as:

$$\mathcal{L}[\mathbf{X}, \boldsymbol{\Omega}] = \mathbf{X}^\dagger \mathbf{A} \mathbf{X} - \mathbf{X}^\dagger \mathbf{X} \boldsymbol{\Omega}$$

Linear problem Lagrangian can be written as:

$$\mathcal{L}[\mathbf{X}] = \mathbf{X}^\dagger \mathbf{A} \mathbf{X} - \mathbf{X}^\dagger \mathbf{P} - \mathbf{P}^\dagger \mathbf{X}$$

Sylvester problem Lagrangian can be written as:

$$\mathcal{L}[\mathbf{X}] = \mathbf{X}^\dagger \mathbf{A} \mathbf{X} - \mathbf{X}^\dagger \mathbf{X} \boldsymbol{\omega} - \mathbf{X}^\dagger \mathbf{P} - \mathbf{P}^\dagger \mathbf{X}$$

These definitions are useful as the Lagrangian is always real even if the matrix elements are not. The Lagrangian is minimized by the desired solution vector, in other words, for the desired solution $\mathbf{X}$:

$$\left.\frac{\delta \mathcal{L}[\mathbf{V}]}{\delta \mathbf{V}}\right|_{\mathbf{V}=\mathbf{X}} = \mathbf{R}(\mathbf{X}) = \hat{\mathbf{A}} \mathbf{X} = \mathbf{0} \qquad \text{(A.1)}$$

The gradient of the Lagrangian, the residual is non-zero for a trial solution or vector that overlaps with the desired analytical solution.

By the Petrov-Galekin condition, $\mathbf{R}^\dagger \mathbf{X} = \mathbf{0}$ for all the above variations, making it useful to attempt to improve the solutions iteratively with $\mathbf{R}$, leading to Krylov subspace methods.

A.2 Note on super matrix notation for RHS

When considering the subtraction of $\mathbf{P}$, components can be subtracted from within the subspace and outside the subspace. As described, this is a form of translation.

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{(V)} \\ \mathbf{P}_{(J)} \end{pmatrix} \tag{A.2}$$

This does not affect the formalism of $\hat{\mathbf{A}}$ for computing residuals, and preconditioning seeks to approximate $\acute{\mathbf{A}}$ which does not include the RHS.

Appendix B

Tully Surface Hopping

B.1 Tully Surface Hopping Algorithm

The Tully Surface Hopping Algorithm as implemented in a custom version of Turbomole 7.6 is listed below. While there are similarities between every step, there are special features distinguishing the first step of the algorithm, and the algorithm with and with out hopping.

Important features of the algorithm are discussed in the main text.

B.1.1 First time step of algorithm

For the first time step of the iteration, $\mathbf{x}(t_0)$ and $\mathbf{v}(t_0)$ are provided, and the density matrix $\boldsymbol{\rho}(t_0)$ is initialized as a zero matrix except the population of the active state; for initial active state k, $\rho_{kk}(t_0) = 1$.

The first time step of the algorithm is hence slightly different:

1. Calculate BO electronic structure properties.

$$\mathbf{V}(t_0) = \mathbf{V}(\mathbf{x}(t_0)), \mathbf{F}(t_0) = \mathbf{F}(\mathbf{x}(t_0)), \mathbf{d}(t_0) = \mathbf{d}(\mathbf{x}(t_0)) \tag{B.1}$$

2. Propagate classical nuclei velocities and positions for half a time step.

$$\mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right) = \mathbf{v}\,(t_0) + \frac{1}{2\mathbf{m}}\mathbf{F}_k(t_0)\Delta t \tag{B.2}$$

$$\mathbf{x}(t_1) = \mathbf{x}(t_0) + \mathbf{v}\left(t_0 + \frac{\Delta t}{2}\right)\Delta t \tag{B.3}$$

3. Construct model Hamiltonian, propagator, and propagate quantum density by half a time step

$$\mathbf{H}(t_0) = \mathbf{V}(t_0) - i\mathbf{v}(t_0) \cdot \mathbf{d}(t_0) \tag{B.4}$$

$$\mathbf{U}(t_0) = \exp(-i\mathbf{H}(t_0)\frac{\Delta t}{2}) \tag{B.5}$$

$$\boldsymbol{\rho}(t_0 + \frac{\Delta t}{2}) = \mathbf{U}(t_0)\,\boldsymbol{\rho}(t_0)\,\mathbf{U}(t_0) \tag{B.6}$$

4. Decisions to attempt hop on next time step:

   - If ($k = 1$ and $\epsilon_1 <$ thres), force a hop to state 0 on next time step and reorder density.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10}, \rho_{0s} \Leftrightarrow \rho_{1s}, \rho_{s0} \Leftrightarrow \rho_{s1} \tag{B.7}$$

B.1.2 Subsequent steps without hopping

The algorithm if a hop is not to occur, with active state k is:

1. Calculate BO electronic structure properties.

$$\mathbf{V}(t_i) = \mathbf{V}(\mathbf{x}(t_i)), \mathbf{F}(t_i) = \mathbf{F}(\mathbf{x}(t_i)), \mathbf{d}(t_i) = \mathbf{d}(\mathbf{x}(t_i)) \tag{B.8}$$

2. Propagate classical nuclei velocities and positions for a time step, calculate average velocity

$$\mathbf{v}\left(t_i + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t_{i-1} + \frac{\Delta t}{2}\right) + \frac{1}{\mathbf{m}}\mathbf{F}_k(t_i)\Delta t \tag{B.9}$$

$$\mathbf{x}(t_{i+1}) = \mathbf{x}(t_i) + \mathbf{v}\left(t_i + \frac{\Delta t}{2}\right)\Delta t \tag{B.10}$$

$$\mathbf{v}_{avg}(t_i) = \frac{1}{2}\left(\mathbf{v}\left(t_i - \frac{\Delta t}{2}\right) + \mathbf{v}\left(t_i + \frac{\Delta t}{2}\right)\right) \tag{B.11}$$

3. Construct model Hamiltonian, propagator and propagate half a time step

$$\mathbf{H}(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_{avg}(t_i) \cdot \mathbf{d}(t_i) \tag{B.12}$$

$$\mathbf{U}(t_i) = \exp(-i\mathbf{H}(t_i)\frac{\Delta t}{2}) \tag{B.13}$$

$$\boldsymbol{\rho}(t_i) = \mathbf{U}(t_i)\,\boldsymbol{\rho}(t_i - \frac{\Delta t}{2})\,\mathbf{U}^\dagger(t_i) \tag{B.14}$$

4. Compute Tully rates and hopping probabilities

$$b_{jk}(t_i) = 2\,\mathrm{Real}(\rho_{jk}(t_i)\mathbf{v}_{avg}(t_i) \cdot \mathbf{d}_{jk}(t_i)) \tag{B.15}$$

$$g_{kj}(t_i) = \frac{b_{jk}(t_i)}{\rho_{kk}(t_i)}\Delta t \tag{B.16}$$

5. Propagate quantum density by another half step

$$\boldsymbol{\rho}(t_i + \frac{\Delta t}{2}) = \mathbf{U}(t_i)\,\boldsymbol{\rho}(t_i)\,\mathbf{U}^\dagger(t_i) \tag{B.17}$$

6. Decisions to attempt hop on next time step:

- Loop from smallest j to largest. If $(\sum_s^j \max(g_{ks}(t_i), 0) > \xi)$, attempt to hop to state j on next time step, exit loop.
- If $(k = 1$ and $\epsilon_1 <$ thres), force a hop to state 0 on next time step and reorder density.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10}, \rho_{0s} \Leftrightarrow \rho_{1s}, \rho_{s0} \Leftrightarrow \rho_{s1} \tag{B.18}$$

The max function carries out Tully's instruction to set negative g_{ks} to zero.

B.1.3 Subsequent step with hopping - velocity rescaling

If a hop from state k to state l is to be tried:

1. Calculate BO electronic structure properties.

$$\mathbf{V}(t_i) = \mathbf{V}(\mathbf{x}(t_i)), \mathbf{F}(t_i) = \mathbf{F}(\mathbf{x}(t_i)), \mathbf{d}(t_i) = \mathbf{d}(\mathbf{x}(t_i)) \tag{B.19}$$

2. Propagate classical nuclei velocity for a half a time step, to obtain old velocity corresponding to state k

$$\mathbf{v}_k\left(t_i\right) = \mathbf{v}\left(t_{i-1} + \frac{\Delta t}{2}\right) + \frac{1}{2\mathbf{m}}\mathbf{F}_k(t_i)\Delta t \tag{B.20}$$

3. Determine velocity rescale factor γ by solving

$$\frac{1}{2\mathbf{m}}\mathbf{d}_{lk}(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma^2 + \mathbf{v}_k(t_i) \cdot \mathbf{d}_{lk}(t_i)\gamma + \epsilon_l(t_i) - \epsilon_k(t_i) = 0 \tag{B.21}$$

4. Decision to hop. The active state is n

- If $\gamma_{\pm}$ is real, select γ_{min} with the smallest magnitude and use it to rescale velocities, to obtain a new velocity corresponding to state l: active state from t_i will be state l

$$\mathbf{v}_n(t_i) = \mathbf{v}_k(t_i) + \gamma_{\min}\frac{\mathbf{d}_{lk}(t_i)}{\mathbf{m}}, n = l \tag{B.22}$$

- Else if $\gamma_{\pm}$ is complex, hop is forbidden, active state remains k

$$\mathbf{v}_n(t_i) = \mathbf{v}_k(t_i), n = k \tag{B.23}$$

5. Propagate new classical velocity by half a time step, propagate classical positions by a full step.

$$\mathbf{v}\left(t_i + \frac{\Delta t}{2}\right) = \mathbf{v}_n\left(t_i\right) + \frac{1}{2\mathbf{m}}\mathbf{F}_n(t_i)\Delta t \tag{B.24}$$

$$\mathbf{x}(t_{i+1}) = \mathbf{x}(t_i) + \mathbf{v}\left(t_i + \frac{\Delta t}{2}\right)\Delta t \tag{B.25}$$

6. Construct model Hamiltonians and propagators with relevant classical nuclei velocity. Propagate quantum density by half a step using old propagator.

$$\mathbf{H}_k(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_k(t_i) \cdot \mathbf{d}(t_i) \tag{B.26}$$

$$\mathbf{U}_k(t_i) = \exp(-i\mathbf{H}_k(t_i)\frac{\Delta t}{2}) \tag{B.27}$$

$$\mathbf{H}_n(t_i) = \mathbf{V}(t_i) - i\mathbf{v}_n(t_i) \cdot \mathbf{d}(t_i) \tag{B.28}$$

$$\mathbf{U}_n(t_i) = \exp(-i\mathbf{H}_n(t_i)\frac{\Delta t}{2}) \tag{B.29}$$

$$\boldsymbol{\rho}(t_i) = \mathbf{U}_k(t_i)\,\boldsymbol{\rho}(t_i - \frac{\Delta t}{2})\,\mathbf{U}_k^{\dagger}(t_i) \tag{B.30}$$

7. Compute Tully rates and hopping probabilities

$$b_{jn}(t_i) = 2\,\mathrm{Real}(\rho_{jn}(t_i)\mathbf{v}_n(t_i) \cdot \mathbf{d}_{jn}(t_i)) \tag{B.31}$$

$$g_{nj} = \frac{b_{jn}(t_i)}{\rho_{nn}(t_i)}\Delta t \tag{B.32}$$

8. Propagate quantum density by another half step using new propagator

$$\boldsymbol{\rho}(t_i + \frac{\Delta t}{2}) = \mathbf{U}_n(t_i)\,\boldsymbol{\rho}(t_i)\,\mathbf{U}_n^{\dagger}(t_i) \tag{B.33}$$

9. Decisions to attempt hop on next time step:

   - Loop from smallest j to largest. If $(\sum_s^j \max(g_{ns}(t_i), 0) > \xi)$, attempt to hop to state j on next time step, exit loop.
   - If ($n = 1$ and $\epsilon_1 <$ thres), force a hop to state 0 on next time step and reorder density.

$$\rho_{00} \Leftrightarrow \rho_{11}, \rho_{01} \Leftrightarrow \rho_{10}, \rho_{0s} \Leftrightarrow \rho_{1s}, \rho_{s0} \Leftrightarrow \rho_{s1} \tag{B.34}$$

B.2 Tully Surface Hopping Trajectory Events

What follows are tables of all the hopping events for a trajectory for Bicyclo[6.2.0]decapentaene molecule at 300K. These tables thus demonstrates the severity of the forbidden hops; however, as discussed in the main text, the cumulative hopping criteria may not be accurate due to a lack of normalization.

Traj.	Event	Cycle	Source	Target	g_{kj}
1	hop	652	2	1	0.03277465638000
	hop	728	1	0	0.00335848831400
	forbidden hop	731	0	1	0.09633590319000
	forbidden hop	743	0	2	0.16857453960000
	forbidden hop	752	0	1	0.11197682600000
	forbidden hop	753	0	1	0.12993629330000
	forbidden hop	753	0	1	0.12993629330000
	forbidden hop	781	0	2	0.13507470430000
	forbidden hop	822	0	2	0.11601194480000
	forbidden hop	906	0	2	0.05234551976000
	forbidden hop	933	0	1	0.20357043230000
	forbidden hop	936	0	1	0.24134377950000
	forbidden hop	956	0	1	0.32938415160000
	forbidden hop	957	0	1	0.06534737903000
	forbidden hop	960	0	2	0.24282198800000
	forbidden hop	968	0	2	0.05404380111000
	forbidden hop	973	0	2	0.21711623440000
	forbidden hop	981	0	2	0.24689419530000
	forbidden hop	987	0	2	0.14084075150000
	forbidden hop	1016	0	2	0.09255572097000
	forbidden hop	1058	0	1	0.10565851380000
	forbidden hop	1066	0	2	0.29454671200000
	forbidden hop	1070	0	1	0.10396096980000
	forbidden hop	1080	0	1	0.03993082955000
	forbidden hop	1085	0	2	0.25041377410000
	forbidden hop	1123	0	2	0.17736489410000
	forbidden hop	1131	0	1	0.06390404925000
	forbidden hop	1151	0	2	0.12154424680000
	forbidden hop	1189	0	1	0.06613465323000
	forbidden hop	1226	0	2	0.15625098280000
	forbidden hop	1253	0	2	0.09262150530000
	forbidden hop	1256	0	2	0.06424530719000
	forbidden hop	1258	0	2	0.05719859533000
	forbidden hop	1259	0	2	0.01740546339000
	forbidden hop	1266	0	1	0.14406882450000
	forbidden hop	1283	0	2	0.12000609560000
	forbidden hop	1284	0	1	0.26260099770000
	forbidden hop	1291	0	2	0.23091262210000
	forbidden hop	1296	0	2	0.20781218430000

Table B.1: All events for selected Bicyclo[6.2.0]decapentaene trajectories at 300K, 1 , labelled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Traj.	Event	Cycle	Source	Target	g_{kj}
1	forbidden hop	1304	0	1	0.06086580400000
	forbidden hop	1318	0	2	0.33235234480000
	forbidden hop	1324	0	2	0.10629914060000
	forbidden hop	1326	0	2	0.20254157390000
	forbidden hop	1349	0	2	0.25757069440000
	forbidden hop	1351	0	1	0.03342332233000
	forbidden hop	1354	0	1	0.03342332233000
	forbidden hop	1385	0	2	0.20452019340000
	forbidden hop	1387	0	2	0.07692552731000
	forbidden hop	1395	0	1	0.03805182758000
	forbidden hop	1421	0	2	0.08731759005000
	forbidden hop	1427	0	2	0.14690542260000
	forbidden hop	1432	0	1	0.10035749930000
	forbidden hop	1433	0	1	0.05192686029000
	forbidden hop	1450	0	1	0.09286082973000
	forbidden hop	1458	0	1	0.03435134576000
	forbidden hop	1461	0	2	0.04429385841000
	forbidden hop	1463	0	2	0.20094187700000
	forbidden hop	1470	0	2	0.01842056120000
	forbidden hop	1481	0	1	0.03509935944000
	forbidden hop	1482	0	2	0.42047453930000
	forbidden hop	1490	0	2	0.21621482660000
	forbidden hop	1495	0	1	0.13801687120000
	forbidden hop	1510	0	1	0.13322310000000
	forbidden hop	1514	0	1	0.23184121410000
	forbidden hop	1518	0	1	0.10895542900000
	forbidden hop	1553	0	1	0.09520982802000
	forbidden hop	1633	0	1	0.08288613503000
	forbidden hop	1667	0	2	0.15167853820000
	forbidden hop	1688	0	2	0.09325535289000
	forbidden hop	1697	0	2	0.03947815745000
	forbidden hop	1725	0	2	0.03082181897000
	forbidden hop	1789	0	1	0.03922776009000
	forbidden hop	1796	0	1	0.12566708310000
	forbidden hop	1818	0	2	0.09711839372000
	forbidden hop	1828	0	1	0.13107239830000
	forbidden hop	1844	0	2	0.06238440580000
	forbidden hop	1847	0	1	0.15253848680000
	forbidden hop	1853	0	1	0.13671403930000

Table B.2: All events for selected Bicyclo[6.2.0]decapentaene trajectories at 300K, 2
, labelled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Traj.	Event	Cycle	Source	Target	g_{kj}
1	forbidden hop	1856	0	1	0.11424521990000
	forbidden hop	1895	0	2	0.31624527570000
	forbidden hop	1900	0	2	0.32089991530000
	forbidden hop	1903	0	1	0.11082565630000
	forbidden hop	1904	0	1	0.03582731232000
	forbidden hop	1917	0	2	0.21410031360000
	forbidden hop	1921	0	1	0.13651638850000
	forbidden hop	1922	0	2	0.68681804670000
	forbidden hop	1925	0	2	0.74877401230000
	forbidden hop	1928	0	1	0.07435141542000
	forbidden hop	1930	0	2	0.40456803670000
	forbidden hop	1934	0	1	0.02920309747000
	forbidden hop	1938	0	2	0.07862288623000
	forbidden hop	1943	0	2	0.05122186578000
	forbidden hop	1951	0	2	0.24920211700000
	forbidden hop	1956	0	2	0.61292756090000
	forbidden hop	1959	0	2	0.53140244960000
	forbidden hop	1962	0	2	0.45538703420000
	forbidden hop	1974	0	1	0.26273433470000
	forbidden hop	1980	0	1	0.37673919250000
	forbidden hop	1983	0	1	0.22049893940000

Table B.3: All events for selected Bicyclo[6.2.0]decapentaene trajectories at 300K, 3 , labelled by the cycle the event is carried out at, with the states involved. The probability of that event is also listed.

Appendix C

Characterization of Electronic Excitations

This chapter contains verbatim excerpts from the following paper. Reprinted (adapted) with permission from *Tetramethylcyclopentadienyl Ligands Allow Isolation of Ln(II) Ions across the Lanthanide Series in* $[K(2.2.2-cryptand)][(C_5Me_4H)_3Ln]$ *Complexes*, Tener F. Jenkins, David H. Woen, Luke N. Mohanam, Joseph W. Ziller, Filipp Furche, and William J. Evans **Organometallics** 2018, 37 (21), 3863-3873 DOI: 10.1021/acs.organomet.8b00557 . Copyright 2018 American Chemical Society.

C.1 Rydberg-like excitations in complexes

For the $4f^n5d^1$ organometallic complexes studied in chapter 4, the Rydberg excitations were characterized for the TDDFT calculation with the TPSSh functional with f-in-core SCecp-mwb-d ECP and basis set on the metal, and def2-SVPD basis set on the ligands. The dominant occupied orbital for these excitation is the αHOMO orbital which has significant

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{La})^-$	918.1(826.4)	0.134	αHOMO+2, 99.0%	0.60	-1.12
	784.3(716.4)	0.078	αHOMO+3, 99.9%	0.74	
	783.8(716.0)	0.078	αHOMO+4, 99.9%	0.74	
$(\mathrm{Cp}'_3\mathrm{La})^-$	654.4(606.4)	0.052	αHOMO+3, 65.5% αHOMO+6, 24.6%	0.76 0.95	-1.42
	633.5(588.4)	0.010	αHOMO+4, 81.2% αHOMO+3, 11.0%	0.83 0.76	
	627.7(583.4)	0.004	αHOMO+5, 88.7% αHOMO+4, 5.1%	0.85 0.83	

Table C.1: Characterization of Rydberg-like excitations in La complexes , with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

metal $5d$ character; the dominant virtual orbital for these excitations have metal $6p$ character, with the lowest energy being metal $6p_z$ in character. The data is listed in tables C.1 through C.6.

C.2 MLCT excitations in complexes

For the $4f^n5d^1$ organometallic complexes studied in chapter 4, the MLCT excitations were likewise characterized for the TDDFT calculation with the TPSSh functional with f-in-core SCecp-mwb-d ECP and basis set on the metal, and def2-SVPD basis set on the ligands. The dominant occupied orbital for these excitation is again the αHOMO orbital which has

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Ce})^-$	893.2(806.1)	0.131	αHOMO+2, 99.0%	0.64	-1.11
	774.9(708.5)	0.079	αHOMO+3, 99.9%	0.77	
	774.3(708.0)	0.078	αHOMO+4, 99.9%	0.77	
$(\mathrm{Cp}'_3\mathrm{Ce})^-$	646.3(599.4)	0.049	αHOMO+3, 67.6% αHOMO+6, 25.0%	0.78 0.96	-1.41
	629.0(584.5)	0.009	αHOMO+4, 79.5% αHOMO+3, 9.5% αHOMO+5, 6.1%	0.85 0.78 0.87	
	623.3(579.6)	0.004	αHOMO+5, 87.0% αHOMO+4, 7.0%	0.87 0.85	

Table C.2: Characterization of Rydberg-like excitations in Ce complexes , with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Pr})^-$	880.2(795.5)	0.128	αHOMO+2, 99.0%	0.67	-1.10
	773.1(707.0)	0.081	αHOMO+3, 99.9%	0.78	
	772.5(706.4)	0.081	αHOMO+4, 99.9%	0.78	
$(\mathrm{Cp}_3'\mathrm{Pr})^-$	639.6(593.6)	0.048	αHOMO+3, 61.6% αHOMO+6, 29.1%	0.79 0.97	-1.43
	621.2(577.8)	0.011	αHOMO+4, 80.5% αHOMO+3, 11.5%	0.86 0.79	
	615.2(572.6)	0.004	αHOMO+5, 89.3% αHOMO+2, 4.6%	0.88 0.75	

Table C.3: Characterization of Rydberg-like excitations in Pr complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Nd})^-$	876.0(792.1)	0.125	αHOMO+2, 99.1%	0.67	-1.10
	777.9(711.0)	0.083	αHOMO+3, 98.6%	0.78	
	777.7(710.8)	0.082	αHOMO+4, 98.6%	0.78	
$(\mathrm{Cp}_3'\mathrm{Nd})^-$	639.6(593.6)	0.045	αHOMO+3, 57.9% αHOMO+6, 28.6% αHOMO+2, 6.5%	0.80 0.99 0.76	-1.42
	623.1(579.4)	0.011	αHOMO+4, 78.7% αHOMO+3, 15.3%	0.87 0.80	
	616.7(573.9)	0.004	αHOMO+5, 91.2%	0.88	

Table C.4: Characterization of Rydberg-like excitations in Nd complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Gd})^-$	994.1(887.4)	0.103	αHOMO+2, 99.6%	0.49	-1.07
	912.2(821.5)	0.083	αHOMO+3, 99.9%	0.57	
	912.0(821.4)	0.082	αHOMO+4, 99.9%	0.57	
$(\mathrm{Cp}_3'\mathrm{Gd})^-$	738.5(677.9)	0.025	αHOMO+2, 70.5% αHOMO+6, 15.9% αHOMO+5, 8.0%	0.63 1.04 0.87	-1.34
	699.1(644.6)	0.005	αHOMO+3, 76.3% αHOMO+5, 12.9% αHOMO+4, 8.5%	0.69 0.87 0.72	
	694.9(641.0)	0.003	αHOMO+4, 66.1% αHOMO+6, 18.2% αHOMO+2, 11.2%	0.72 1.04 0.63	

Table C.5: Characterization of Rydberg-like excitations in Gd complexes , with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Tb})^-$	1101.9(972.3)	0.092	αHOMO+2, 99.8%	0.37	-1.05
	1015.0(904.0)	0.079	αHOMO+3, 99.9%	0.45	
	1014.6(903.7)	0.078	αHOMO+4, 99.9%	0.45	
$(\mathrm{Cp}_3'\mathrm{Tb})^-$	800.5(729.8)	0.022	αHOMO+2, 91.1%	0.50	-1.33
	745.8(684.1)	0.002	αHOMO+3, 41.4% αHOMO+5, 36.6% αHOMO+4, 21.2%	0.57 0.88 0.62	
	743.3(681.9)	0.003	αHOMO+4, 77.6% αHOMO+3, 10.8% αHOMO+5, 9.9%	0.62 0.57 0.88	

Table C.6: Characterization of Rydberg-like excitations in Tb complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

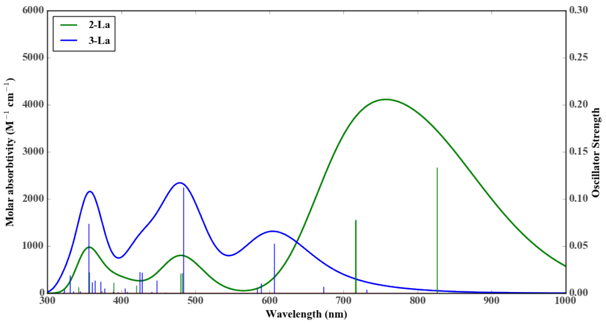

Figure C.1: Predicted UV-vis absorption spectra for 2-La and 3-La

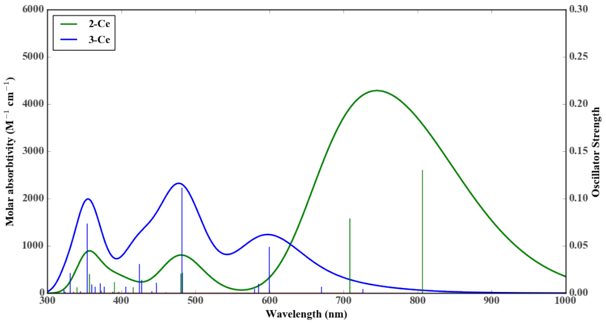

Figure C.2: Predicted UV-vis absorption spectra for 2-Ce and 3-Ce

significant metal 5d character; the dominant virtual orbital has significant π^* character on the ligands. The data is listed in tables C.7 through C.12.

C.3 f-out-of-core calculations for 2-La and 3-La

As described in the main text, these calculations are possible for La, are reported in Table C.13, and can be compared to Table C.1 and C.7.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{La})^-$	511.6(481.8)	0.021	αHOMO+8, 92.4%	1.63	-1.12
	509.6(480.1)	0.020	αHOMO+9, 97.9%	1.64	
$(\mathrm{Cp}'_3\mathrm{La})^-$	513.6(483.6)	0.112	αHOMO+7, 96.6%	1.35	-1.42

Table C.7: Characterization of MLCT excitations in La complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Ce})^-$	511.8(481.9)	0.021	αHOMO+8, 69.8% αHOMO+7, 27.8%	1.64 1.63	-1.11
	509.8(480.2)	0.020	αHOMO+9, 96.9%	1.64	
$(\mathrm{Cp}'_3\mathrm{Ce})^-$	511.5(481.7)	0.112	αHOMO+7, 96.3%	1.36	-1.41

Table C.8: Characterization of MLCT excitations in Ce complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Pr})^-$	511.9(482.1)	0.022	αHOMO+7, 83.5% αHOMO+9, 11.2%	1.64 1.65	-1.10
	510.1(480.4)	0.021	αHOMO+8, 79.1% αHOMO+9, 18.9%	1.65 1.65	
$(\mathrm{Cp}_3'\mathrm{Pr})^-$	506.5(477.3)	0.111	αHOMO+7, 96.2%	1.37	-1.43

Table C.9: Characterization of MLCT excitations in Pr complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Nd})^-$	511.0(481.3)	0.021	αHOMO+7, 94.8%	1.65	-1.10
	510.7(480.9)	0.021	αHOMO+8, 93.8%	1.65	
$(\mathrm{Cp}_3'\mathrm{Nd})^-$	505.0(475.9)	0.111	αHOMO+7, 95.7%	1.38	-1.42

Table C.10: Characterization of MLCT excitations in Nd complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Gd})^-$	514.4(484.3)	0.024	αHOMO+7, 93.4%	1.66	-1.07
	513.7(483.7)	0.024	αHOMO+8, 94.0%	1.66	
$(\mathrm{Cp}_3'\mathrm{Gd})^-$	514.6(484.4)	0.115	αHOMO+7, 61.5% αHOMO+8, 34.1%	1.41 1.42	-1.34

Table C.11: Characterization of MLCT excitations in Gd complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{Tb})^-$	516.6(486.2)	0.027	αHOMO+7, 95.1%	1.66	-1.05
	515.6(485.3)	0.027	αHOMO+8, 96.2%	1.67	
$(\mathrm{Cp}_3'\mathrm{Tb})^-$	517.7(487.2)	0.117	αHOMO+7, 85.9% αHOMO+8, 9.8%	1.41 1.44	-1.33

Table C.12: Characterization of MLCT excitations in Tb complexes
, with the Kohn-Sham orbital energies reported. Wavelengths blue shifted by 0.15eV in parenthesis.

	Wavelength (nm)	Oscillator Strength	Dominant virtual orbital	Dominant virtual orbital energy (eV)	Dominant occupied orbital energy (eV)
$(\mathrm{Cp}_3^{tet}\mathrm{La})^-$	893.9	0.123	αHOMO+1, 98.8%	0.60	-1.17
	742.7	0.051	αHOMO+3, 99.8%	0.77	
	742.6	0.051	αHOMO+4, 99.8%	0.77	
	501.8	0.019	αHOMO+7, 995.1%	1.64	
	501.7	0.019	αHOMO+8, 94.8%	1.64	
$(\mathrm{Cp}_3'\mathrm{La})^-$	641.3	0.054	αHOMO+3, 71.8% αHOMO+6, 21.5%	0.76 0.93	-1.46
	618.1	0.009	αHOMO+4, 87.6% αHOMO+3, 7.1%	0.83 0.76	
	611.0	0.004	αHOMO+5, 90.4%	0.85	
	503.8	0.103	αHOMO+7, 95.5%	1.36	

Table C.13: Characterization of excitations in La complexes
, with the Kohn-Sham orbital energies reported.

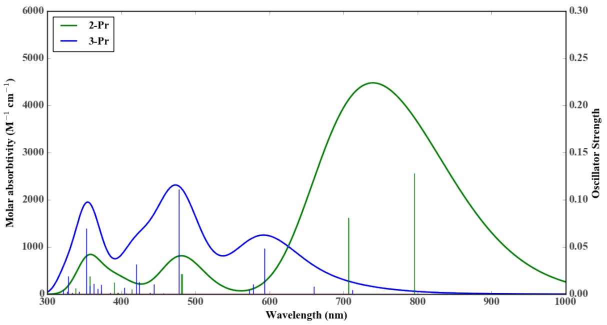

Figure C.3: Predicted UV-vis absorption spectra for 2-Pr and 3-Pr

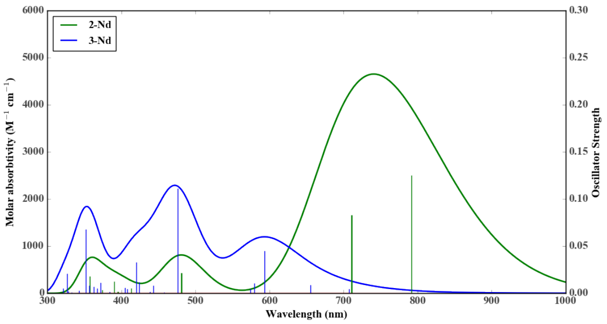

Figure C.4: Predicted UV-vis absorption spectra for 2-Nd and 3-Nd

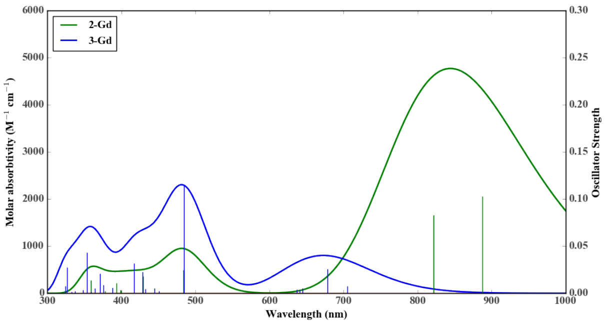

Figure C.5: Predicted UV-vis absorption spectra for 2-Gd and 3-Gd

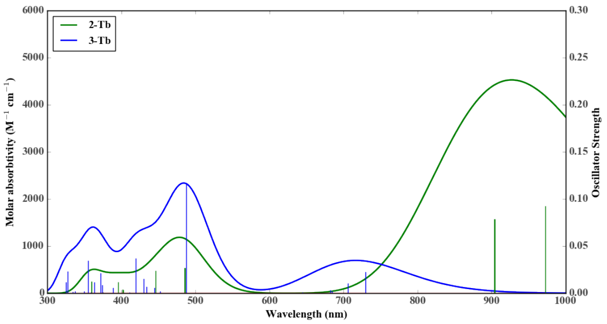

Figure C.6: Predicted UV-vis absorption spectra for 2-Tb and 3-Tb

Appendix D

Resources for Intermolecular Forces game

The Supporting Information is available on the ACS Publications website at DOI:

10.1021/acs.jchemed.0c00050

These include the pre-/postactivity test, a guiding worksheet for the students, and the card game with instructions.

All materials are also provided in the standard Microsoft Word format at the time of publication to allow for direct changes to the activity.