**Title**
Self Consistent Excited State Mean Field Theory: Development and Applications



**Author**
Hardikar, Tarini Shekhar

**Publication Date**
2023

Peer reviewed|Thesis/dissertation

# Self Consistent Excited State Mean Field Theory: Development and Applications

By

Tarini S. Hardikar

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eric Neuscamman, Chair
Professor David Limmer
Professor Lin Lin

Spring 2023

Self Consistent Excited State Mean Field Theory: Development and Applications

Abstract

Self Consistent Excited State Mean Field Theory: Development and Applications

by

Tarini S. Hardikar

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Eric Neuscamman, Chair

In the wide spectrum of excited state quantum chemistry methods, there is no direct analogue to Hartree-Fock theory. This dissertation presents the theory and initial applications for a self consistent framework for Excited State Mean Field (ESMF) theory. This method presents a self consistent equation analogous to the Roothaan-Hall equation, that is constructed with mean-field one-electron operators. The convergence criteria is described by a commutator condition between Fock-type operators and density operators, just like in Hartree-Fock theory. Finally, this method is accelerated via direct inversion of the iterative subspace (DIIS), akin to acceleration in the ground state theory. Futhermore, this work discusses applications of ESMF to larger solvated systems, afforded by ESMF's scaling – the method costs roughly twice the cost of a Hartree-Fock calculation. Getting an accurate physical picture of excited states in solvated systems is challenging, and the second half of this dissertation focuses on a comparative analysis of various methods, their degree of correlation, and their ability to qualitatively describe donor and acceptor regions for a charge transfer excitation. This comparison shows that ESMF can accurately describe the degree of participation of solvent water molecules in the excitation, unlike Density Functional Theory (DFT) based methods. However, when there is minimal participation from the solvent, Restricted Open-shell Kohn Sham methods fare better, indicating that the lack of correlation in ESMF prevents the method from providing a more quantitatively accurate picture. Using ESMF as a stepping rung for developing a hierarchy of excited state specific methods is a promising platform to achieve affordable excited state specific calculations.

***To my grandparents:***

My maternal grandfather, *nana*, who I never got to meet;

My maternal grandmother, *nani*, who I associate with relaxed summer vacations, warmth, and hospitality;

My paternal grandfather, *ajoba*, who always wanted me to get a PhD, who passed away too soon;

My paternal grandmother, *aaji*, who's always prioritized education, and who's asked me if I have studied and eaten, every week for the last ten years.

# Contents

**Contents** **ii**

**List of Figures** **iv**

**List of Tables** **v**

**1 Introduction** **1**
1.1 Quantum Chemistry Fundamentals . . . . . 1
1.2 Born-Oppenheimer Approximation . . . . . 2
1.3 Types of Excitations . . . . . 3
1.4 Hartree-Fock Method . . . . . 4
1.5 Direct optimization methods . . . . . 6
1.6 Electronic Correlation . . . . . 7
1.7 Ground State Methods . . . . . 7
1.8 Excited State Methods . . . . . 9
1.9 Outlook . . . . . 14
1.10 Outline . . . . . 14

**2 A Self Consistent Field Formulation of Excited State Mean Field Theory** **16**
2.1 Abstract . . . . . 16
2.2 Introduction . . . . . 16
2.3 Theory . . . . . 18
2.4 Results . . . . . 23
2.5 Conclusion . . . . . 26

**3 Exploring Mulliken Population Changes upon Excitation in Explicitly Solvated Systems** **27**
3.1 Abstract . . . . . 27
3.2 Introduction . . . . . 27
3.3 Theory . . . . . 29
3.4 Computational Details . . . . . 32
3.5 Results . . . . . 32

3.6 Conclusion . . . . . 40

**4 Conclusion** **41**

**Bibliography** **42**

**5 Appendix 1: A Self Consistent Field Formulation of Excited State Mean Field Theory** **56**
5.1 Mathematical Detail . . . . . 56
5.2 Charge Transfer Iteration Details . . . . . 63
5.3 Geometries . . . . . 63

**6 Appendix 2: Exploring Mulliken Population Changes upon Excitation in Explicitly Solvated Systems** **66**
6.1 Geometries . . . . . 66

# List of Figures

1.1 A graphical depiction of vertical versus adiabatic excitation energies where $R$ is the distance coordinate. $\Delta E_{vertical}$ refers to vertical excitation energies, $\Delta E_{adiabatic}$ refers to adiabatic excitation energies. Image adapted from [3]. . . . . 3
1.2 Graphical description of the $\Delta$-SCF and ROKS orbital relaxation process. The former requires two separate optimizations, followed by a spin purification. ROKS requires only one orbital optimization. Image adapted from [64]. . . . . 11
1.3 Relationship between various ground state and response state methods. Figure adapted from [13]. . . . . 13

2.1 Donor (a) and acceptor (b) orbitals for the lowest charge transfer state in the PYCM molecule as predicted by ESMF. . . . . 25

3.1 Lewis structures for formaldehyde (left), 4-cyanopyridine (center), and PYCM (right). . . . . 31
3.2 The hole (at left) and particle (at right) orbitals for the back-bonding to $\pi^*$ excitation in formaldehyde with 12 waters as calculated by ESMF. . . . . 34
3.3 The hole (at left) and particle (at right) orbitals for the back-bonding to $\pi^*$ excitation in formaldehyde with 12 waters as calculated by ROKS with the $\omega$B97X-V functional. . . . . 34
3.4 Participating HOMO and LUMO orbitals for ROKS and TDDFT for the $\omega$B97X-V functional for 4-cyanopyridine with a 4 water solvation shell. Top left: HOMO calculated by ROKS. Top right: LUMO calculated by ROKS. Bottom left: HOMO calculated by TDDFT. Bottom right: LUMO calculated by TDDFT. . . . . 36
3.5 Relevant HF orbitals in PYCM with 10 water molecules. . . . . 37
3.6 Comparison of hole and particle orbitals in PYCM with 10 water molecules. . . 39

# List of Tables

2.1 Convergence of SCF- and GVP-based ESMF for the HOMO/LUMO excitation of cc-pVDZ $H_2O$ . . . . . 23
2.2 Total time in seconds and number of iterations $n_i$ taken for the orbital optimization in the ground state (for RHF) or the excited state (for SCF-based ESMF) to get within $5\mu E_h$ of its fully converged value. . . . . 23
2.3 Convergence of the energy for the lowest singlet core excited state of $H_2O$ in the aug-cc-pVDZ basis. . . . . 25

3.1 Changes in the Mulliken charges of different regions for the formaldehyde $n \rightarrow \pi^*$ excitation with 8 and 12 waters. . . . . 33
3.2 Changes in the Mulliken charges of different regions for the formaldehyde back-bonding$\rightarrow \pi^*$ excitation with 12 waters. . . . . 35
3.3 Changes in Mulliken charges during the 4-cyanopyridine $n \rightarrow \pi^*$ excitation in gas phase and with 4 explicit waters. Note that ROKS PBE0 did not converge for the latter case. . . . . 36
3.4 Changes in Mulliken charges for the PYCM CT state with 4, 10, and 20 water molecules. Note that ROKS PBE0 did not converge for the largest case. The Ref. data is EOM-CCSD in the 4 water case and CASSCF in the 10 water case. 38

5.1 ESMF details in PYCM charge transfer state. . . . . 62

## Acknowledgments

In the pre-pandemic days, my office mates and I would keep a quote wall – the most memorable jokes, ridiculous puns, and silly thoughts, all tacked up with sticky notes in a corner. In my last semester as a graduate student, I've thought a lot about one of these quotes - something I said as a first-year, "I hate science today, it's a good day to meet with Eric." Five years later this stays true, and my first thanks are to my advisor, Eric. Through my own ups and downs with science (really with scientific software), Eric has provided his support, his guidance, and his relentless optimism. I don't know anyone else who is so confident that something could always be implemented in a day, when it would almost always take a month. For this perspective and for his mentorship, I'm very grateful.

I would also like to thank my qualifying exam committee: Professors Phill Geissler, Birgitta Whaley, Naomi Ginsberg, and Lin Lin. In their own way, they provided me with new ways to think about my work and present it confidently. I deeply wish I could show Phill that I did learn to think about the chemistry, and not just the math, eventually. I would also like to thank my dissertation committee, Professors David Limmer and Lin Lin for their support and patience.

My gratitude for fellow Neuscamman group members is in four parts: First, for the group alumni: Jacki, Luning, Leon, Beatrice, and Sergio for all your mentorship. I am grateful I had people to look up to, learn from, and debug with. Special thanks to Jacki for helping me get started with research as a first-year. A huge thanks to Sonja and Trine, for their wonderful friendship and encouragement through these last few years. Thank you to Harrison and Brianna, being your formal mentor in the group was a great experience, and has made me a better scientist. And finally, my fellow Neuscamman cohort mates: Rachel, Becky, Scott, and Connie. Special thanks to Rachel for helping run calculations and being a great sounding board; and to Scott, for helping with code implementation, and providing technical fixes. I am so grateful to have had such wonderful company through graduate school milestones. For all the late nights in the basement of Gilman, to all the random Zooms through the pandemic, for all the camaraderie, thank you.

I would be remiss without thanking members of the Martin Head-Gordon group for all their help with QChem, Leo, Abdul, Adam, Lexie, Brad, Kevin, and Jiashu. Thank you for helping me navigate the labyrinths of genscfman. Special thanks to Leo; my work would not be implemented in QChem without his guidance. I would also like to thank Dr. Hugh Burton for helping me with QChem.

I would like to thank both the Chemistry Graduate Life Committe and UAW 2865 for giving me a community. I would also like to thank Matt Francis and Joel Adlen – I've probably emailed them at least once a month my entire time at Berkeley, and truly, they've always managed to fix everything.

I would also like to thank my partner, Aagam, for all his support, encouragement, and strength through these years. I am grateful for all the pep talks, the practical advice, his unwavering presence, his confidence that I could do this, and all the patience in the world as I ranted about ceilings that leak and code that didn't compile, and algorithms that didn't

converge. Finally, I would like to thank my family: my parents Shekhar and Anjali, my brother Tanmay, my aunts Mamata and Neeta, and my grandmother for their love, their presence, their cheerleading, and their consistent faith in my ability to eventually figure it out. I would not be here without you.

To all of these people, my deepest gratitude, thank you.

# Chapter 1

# Introduction

## 1.1 Quantum Chemistry Fundamentals

The fundamental equation governing quantum chemistry is the time dependent Schrödinger equation [1]. This equation determines how a wavefunction propagates in time, and various mathematical properties of this equation explain various physical properties of the system.

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

Of course, this equation assumes relativistic factors don't play a huge role in the study of systems of interest. In practice however, even without relativistic effects, this equation is far too complex and unwieldy to be solved in its entirety [1], and quantum chemists make various approximations to suit their needs and applications. This work is principally concerned with the electronic structure of molecules, that is, how electrons behave in their ground state and upon excitation. To think about the electronic structure theory, this work makes two key simplifications:

The first is time independence. We work with the time independent Schrödinger equation, where the equation is an eigenvalue problem

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

where for a wavefunction, $\psi$, the time-independent Hamiltonian operator, $\hat{H}$ gives the energy of the system as its eigenvalues. The eigenvectors here are stationary states, and are critically important in studying spectroscopy [2]. The time independent equation focuses on solving for these stationary states given their importance. So, this reduces the problem at hand to an eigenvalue equation where the operator is the quantum mechanical Hamiltonian, given as [1]:

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{1.3}$$

where $N$ is the number of electrons in the system, $M$ the number of nuclei, lowercase indices track electrons, uppercase indices track nuclei, $M_A$ is the ratio of mass of nucleus $A$ to the mass of an electron, $Z_A$ is the atomic number of nucleus $A$, and the Laplacian subscripts determine what particle's position the differentiation is with respect to. This equation (and following equations in this work) is in atomic units, where electron charge, electron mass and $\hbar$ are set to 1. The first term represents the kinetic energy operator for electrons, the second term is the kinetic energy of the nuclei, the third term is the Coulomb interaction between electrons and nuclei, the fourth term is the repulsion between electrons, and the fifth term is the repulsion between nuclei.

## 1.2 Born-Oppenheimer Approximation

The second simplification we make is the Born Oppenheimer approximation. In most applications of quantum chemistry, and in all of the research presented in this thesis, quantum chemists employ the Born Oppenheimer approximation, which makes the eigenvalue problem significantly more tractable. This approximation states that since the mass of the nuclei is thousands of times greater than the mass of electrons, the timescale of the motion of nuclei is much slower than that of electrons. Therefore, the nuclei can be taken to be stationary with respect to electronic motion [1, 2]. This reduces the Hamiltonian to the following terms:

$$\hat{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{1.4}$$

Note that now, we are also looking at an electronic wavefunction that depends explicitly on the position of the electrons, but only parametrically on the position of nuclei. The total energy for a fixed geometry of the nuclei is then given by the semi-classical expression

$$\mathcal{E}_{tot} = \mathcal{E}_{elec} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{1.5}$$

where the nuclear kinetic energy is ignored and the nuclear-nuclear potential energy is given by the classical Coulomb energy for point charges. A quantum mechanical Hamiltonian for the nuclei is arrived at by taking an average of the electronic wavefunction to generate a nuclear Hamiltonian.

$$\hat{H}_{nucl} = -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \mathcal{E}_{tot}(\{R_A\}) \tag{1.6}$$

where the total energy provides a potential energy for nuclear motion by considering averaged electronic motion. Thus, the total energy of the system is described by the nuclear kinetic energy, the electronic energy, and the point-charge nuclear repulsion energy. As stated above,

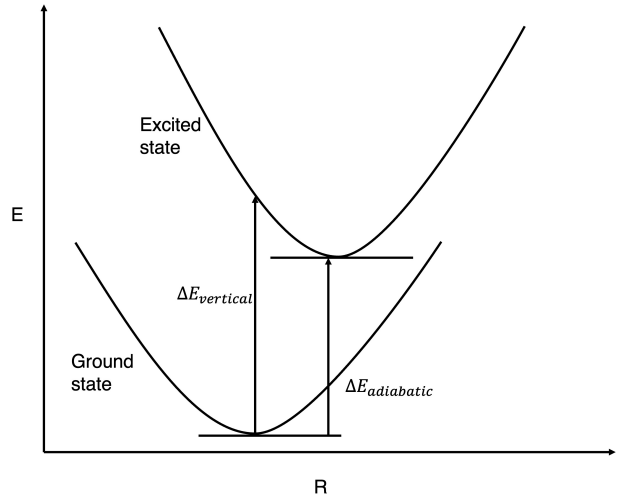


**Figure 1.1:** A graphical depiction of vertical versus adiabatic excitation energies where $R$ is the distance coordinate. $\Delta E_{vertical}$ refers to vertical excitation energies, $\Delta E_{adiabatic}$ refers to adiabatic excitation energies. Image adapted from [3].

given the focus on the electronic structure of the molecule, this work will ignore the nuclear energies and the translational, vibrational, and rotational degrees of freedom will not be discussed. Instead, this work will discuss the electronic energy of the ground and excited states of molecules based on the Hamiltonian in Eq 1.4.

## 1.3 Types of Excitations

Specifically, this work focuses on the vertical excitation energies of molecules. As described in Fig. 1.1, this means that we are looking at the molecule instantaneously after excitation, and not allowing it to relax to a favorable new geometry. Because of this constraint, the vertical excitation energy is typically higher than the relaxed (adiabatic) excitation energy [3].

While there are many types of electronic excitations, this work primarily focuses on two: valence excitations and charge transfer excitations. Valence excitations are when both the orbital the electron is excited out of and the orbital the electron is excited into are valence orbitals and are nearby one another. An example of such as excitation would be an electron moving from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). The second type of excitation that's of interest here is a charge transfer excitation, where the electron moves from one set of orbitals spatially separated set of orbitals [4, 5]. These are marked by a separation of charge, and could be intermolecular or intramolecular charge transfers. While core excitations [6] and Rydberg excitations [7] are briefly discussed, they are not the main subject of attention in this work.

Note that in some molecules low-lying Rydberg excitations lie interspersed between valence excitations [8].

Even after making the Born-Oppenheimer approximation, the electronic problem cannot be solved analytically for multi-electron systems, even for the ground state, and requires approximations [9]. There are a plethora of various methods that rely on various types and levels of approximations. Broadly, methods can be sorted as wavefunction based or density functional based methods. The latter methods fall under the broad umbrella of density functional theory (DFT) and are by far the most commonly used methods in the wider chemistry community [10]. However, density functionals are approximate in practice, and results can be sensitive to the choice of which functional approximation is employed [11, 12]. One the other hand, DFT has the strong advantage of a relatively low computational cost, with modern functionals' costs growing as the third or fourth power of the system size depending on the approximation and implementation details [13–15].

Wavefunction based methods, can offer higher accuracy than DFT but at increased computational cost. In quantum chemistry, most wave function methods are built atop the Hartree-Fock method.

## 1.4 Hartree-Fock Method

This method approximates the electron-electron coulomb interaction as an effective mean-field interaction in which electrons feel a fermionic-antisymmetry-preserving average of the repulsion from other electrons. Thus, the electron-electron interaction is reduced to an averaged cloud felt by all electrons. The wavefunction is approximated as a single Slater determinant, and the minimization follows the self consistent field procedure [1]. A brief summary of the method is as follows. A Slater determinant is defined as

$$\psi(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \ldots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \ldots & \chi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_N) & \chi_2(x_N) & \ldots & \chi_N(x_N) \end{vmatrix} \tag{1.7}$$

where the $\frac{1}{\sqrt{N!}}$ is a normalization factor and the determinant places $N$ electrons with positions $x_1, \ldots, x_N$, in $N$ spin orbitals $\chi_1, \ldots, \chi_N$, without specifying which electron is in which orbital. The determinant properties ensure an overall antisymmetric form. One way to motivate this wave function form is that it is the result of antisymmetrizing a Hartree Product. Minimization of the energy of this Slater determinant with respect to varying the shapes of the orbitals leads to a one-electron eigenvalue problem involving a mean field Fock operator [1]

$$\boldsymbol{F}(d) = h + \sum_b \mathcal{J}_b(d) - \mathcal{K}_b(d) \tag{1.8}$$

Here, $d$ is an $N$-electron density, the first term is the one-electron part of Eq 1.4, and the second summation is over spin orbitals $\chi$, describing the two-electron interactions, split into the classic Coulomb term ($\mathcal{J}$ repulsive) and the strictly quantum "exchange" term ($\mathcal{K}$ attractive).

Then, the Hartree-Fock equation can be represented as an eigenvalue equation

$$\boldsymbol{F}|\psi_a\rangle = \varepsilon_a|\psi_a\rangle \tag{1.9}$$

Approximating the orbitals as a linear combination of basis functions leads to the Roothaan-Hall equation

$$\boldsymbol{FC} = \boldsymbol{SC\varepsilon} \tag{1.10}$$

where $C$ is a matrix of molecular orbital coefficients, $S$ is the overlap matrix between basis functions, and $\varepsilon$ is a diagonal matrix of orbital energies. While the task of minimizing the energy of a Slater determinant with respect to the orbital coefficients $C$ can be thought of as a standard nonlinear minimization problem and can be treated as such [16], it is far more common to find the Hartree-Fock state via a self-consistent procedure. With the self consistent procedure, for a given choice of $C$, the pseudo-eigenvalue equation is solved, a new $C$ is generated, and this procedure is repeated until both sides of the equation match.

A self-consistent procedure is often preferred to a direct minimization, even though it can occasionally lead to variational collapse, wherein a lower eigenstate than desired is found. Methods such as variations on Hartree-Fock theory (Projected Hartree-Fock [17, 18]) as well as DFT methods ([19–21]) are built on a self-consistent procedure. Its efficiency makes it the default optimization choice in common quantum chemistry packages as well [22–24]. A common method to accelerate SCF procedures is the direct inversion of the iterative subspace (DIIS) method [25]. At any given SCF iteration, the residuals $\boldsymbol{\{e\}}$ from previous iterations are calculated and a linear combination is constructed

$$e_{m+1} = \sum_{i}^{m} c_i e_i \tag{1.11}$$

DIIS aims to minimize this residual norm with the constraint that the coefficients ($\boldsymbol{c}$) are normalized. These coefficients are calculated by solving the a Lagrangian equation with the multiplier $\lambda$ given by [25]

$$\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1m} & 1 \\ B_{21} & B_{22} & \dots & B_{2m} & 1 \\ \vdots & \vdots & \ddots & \vdots & 1 \\ B_{m1} & B_{m2} & \dots & B_{mm} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \\ -\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{1.12}$$

The coefficients obtained from solving this equation are then used to update the variable value. DIIS has proven to be an incredible tool in speeding SCF convergence [26], and quantum chemistry packages normally use it as a default accelerator [22].

## 1.5 Direct optimization methods

When the SCF procedure fails to converge, direct optimization methods are needed. These methods are also necessary for theories that might not have a SCF solution. In such cases, methods such as gradient descent, Newton-Raphson, BFGS [27, 28], or Geometric Descent Method [16] might be used. Two recent optimization techniques specific to electronic structure theory are described below.

### 1.5.1 Square gradient minimization

In this method, a extremization problem is converted to a minimization problem by focusing on a slightly different objective function [29]

$$\Delta = |\nabla_\theta \mathcal{L}|^2 = \sum_{ai} \left| \frac{\partial \mathcal{L}}{\partial \theta_{ai}} \right|^2 \tag{1.13}$$

where the Lagrangian is with respect to some orbital constraint described by $\theta$. $\mathcal{L}$, in practice, is the total energy so this approach ultimately aims to zero out the energy gradient by minimizing the square gradient norm. The resulting descent method safeguards against variational collapse, and only requires Lagrangian gradients. However, it costs 3 times as much as ground state optimization per iteration [29]. SGM has been used to converge ROKS and TDDFT calculations, especially for core excitations and X-ray spectroscopy applications [30–32].

### 1.5.2 Generalized Variational Principle

In this method, a generalized form of the variational principle is developed, that in principle works with a wide range of electronic structure methods as long as they can be defined by a list of property vectors [33]. This principle is given by

$$\lim_{\mu \to 0} \min_{\vec{c}} (\mu |\vec{d}|^2 + (1-\mu) |\nabla E|^{\ 2}) \tag{1.14}$$

where $\vec{c}$ is a list of wavefunction variables, and $\vec{d}$ is the list of property vectors (such as Mulliken properties, symmetry, or charge). Here, $\mu$ is a "toggle" function that helps target a specific state without variational collapse. Similar to SGM, GVP also aims to find the zero point of the energy gradient, but by pushing the optimiziation towards the desired excited state via the property vectors. In practice, this method allows the user to be within the region of the solution with the $\nabla E$ term, and then hone in on the state with the property vectors. This method provides a strong foundation on which other excited state specific theories can be built [34–36].

## 1.6 Electronic Correlation

Despite Hatree-Fock's rather rudimentary approximation of the electronic field, this method provides a robust starting point for more advanced methods, and already captures a large fraction of the "true" energy of the system. The difference between the actual eigenvalues of the Hamiltonian and the Hartree-Fock energy is described as the correlation energy [37]. Other wavefunction based methods (often described as post-Hartree-Fock methods) aim to capture this correlation.

The correlation energy is divided into two main phenomena, static (or strong correlation) and dynamic (or weak correlation) [38]. In strongly correlated systems, the underlying assumption from Hartree-Fock theory that the electronic configuration can be described by a single determinant is in itself incorrect. In such systems, this is known as having high multi-reference character, a classic example being degenerate states with equal probability of an $\alpha$ or $\beta$ spin excitation. Dynamic correlation is when the Hartree-Fock electron configuration is qualitatively sufficient, but requires corrections to capture the right quantitative measure of electronic correlation.

## 1.7 Ground State Methods

Since the development of Hartree-Fock theory, there have been hundreds of methods trying to capture different aspects of the correlation energy. The most widely used wavefunction-based correlation methods rely on Hartree-Fock theory as a starting point in which mean-field orbital relaxations have already been applied to the wavefunction [1]. Given the focus of this work, only certain methods will be discussed.

### 1.7.1 Configuration Interaction

A straightforward method to add strong correlation is to consider the Configuration Interaction method where in addition to the Aufbau Slater determinant, determinants with every possible excitation are considered [1]:

$$|\psi_{CI}\rangle = C_0|\psi_0\rangle + \sum_{ia} C_{ia}|\psi_i^a\rangle + \sum_{ijab} C_{ij}^{ab}|\psi_{ij}^{ab}\rangle + \dots \tag{1.15}$$

where the various $C$ reflect the coefficients of the wavefunction components, $i, j$ refer to occupied orbitals, and $a, b$ refer to the virtual orbitals. In other words, the simple Aufbau Slater determinant is supplemented by a weighted sum of every single, double, etc excitation on this determinant. The space of considering all excitations is called Full Configuration Interaction [39, 40]. Of course, this is incredibly computationally expensive, and scales combinatorially with system size [2].

In practice, a truncated wavefunction may be considered, and the method is named accordingly: Configuration Interaction Singles (CIS) for only singles excitations, Configuration

Interaction Singles Doubles (CISD) for single and double excitations, and so on. CIS scales as $\mathcal{O}(N^4)$ and CISD as $\mathcal{O}(N^6)$ where $N$ is the number of one-electron basis functions [2, 13]. Despite reasonable scaling, these truncated CI wavefunctions are not very compact, and tend to converge slowly [41]. Additionally, CISD is not size consistent and therefore becomes less accurate as system size increases [2].

### 1.7.2 Perturbation Theory

One of the most commonly used methods to add weak correlation to the Hartree-Fock method is with second order perturbation theory, commonly referred to as the Möller Plesset perturbation theory [2]. A perturbed Hamiltonian is considered as

$$\hat{H} = \boldsymbol{F} + \hat{V} \tag{1.16}$$

where $\hat{V}$ is the perturbative field accounting for correlation effects neglected in Hartree-Fock theory. The second order energy correction is give by [1]

$$E_{MP2} = \sum_{ijab} \frac{|\langle ab||ij\rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{1.17}$$

where the numerator is $\langle ab||ij\rangle = \langle ab|ij\rangle - \langle ab|ji\rangle$ in physicist notation, and $\varepsilon$ are the Fock operator eigenvalues from Eq 1.10. Physicists' notation is a common shorthand for expressing two-electron interactions, where for spin orbitals $i, j, k, l$ the notation is as follows [1]

$$\langle ij|kl\rangle = \int \mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2}\chi_i^*(\boldsymbol{x_1})\chi_j^*(\boldsymbol{x_2})\frac{1}{r_{12}}\chi_k(\boldsymbol{x_1})\chi_l(\boldsymbol{x_2}) \tag{1.18}$$

Of course, higher order perturbative theories can also be used, though this second order method is used the most commonly. MP2, as it is colloquially referred to, scales as order $\mathcal{O}(N^5)$ where $N$ referes to the number of basis functions [1]. Additionally, it also has the advantage of being size consistent, and therefore maintains accuracy as system size increases [41, 42].

### 1.7.3 Coupled Cluster

A more sophisticated approach to wavefunction-based many-body perturbation theory is the coupled cluster (CC) method [43, 44]. This method, unlike CI, is size extensive, and therefore maintains its accuracy as system size increases [45]. In CC methods, the exponential of an excitation operator $\hat{T}$, called a cluster operator, is applied to the Hartree-Fock ground state wavefunction.

$$|\psi_{CC}\rangle = \exp(\hat{T})|\psi_0\rangle \tag{1.19}$$

The cluster operator is broken down as

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

where $\hat{T}_1$ is the operator for single excitations, $\hat{T}_2$ is the operator for double excitations and so on. The general form of the cluster operator, in second quantized algebra, is given as

$$\hat{T}_n = \frac{1}{(n!)^2} \sum_{ijkpqr\ldots} t_{ijk\ldots}^{pqr\ldots} p^\dagger i q^\dagger j r^\dagger k \ldots \tag{1.21}$$

where $i, j, k, \ldots$ are the occupied orbitals, $p, q, r, \ldots$ are the virtual orbitals [44]. So essentially, the cluster operator is similar to a term in the CI wavefunction - a weighted sum of excitations of each order. Then, CC inserts the exponential ansatz into the time-independent Schrödinger equation [2]

$$\hat{H} \exp(\hat{T})|\psi_0\rangle = E \exp(\hat{T})|\psi_0\rangle \tag{1.22}$$

and then takes left projections with different Slater determinants to create a system of equations for the energy and amplitudes.

Again, just as in CI, CC methods are named as CCS, CCSD, and so on, based on the order of the excitations considered. CCSD scales as order of $\mathcal{O}(N^6)$, where again, $N$ is the number of basis functions [2].

## 1.8 Excited State Methods

Given the approximations and challenges involved with ground state electronic structure theory, it is no surprise that excited state electronic structure theory is more complicated, and requires both new theory, and more nuance. While there is typically only one global ground state, excited states are plentiful, with different spin and symmetry character, and they are more difficult to optimize as they are not energy minima but instead typically exist as saddle points [46]. One of the most straightforward approaches again, is to look at higher roots of the CI equations. Depending on what order of excitations one includes in the CI, this can, as for the ground state, be prohibitively expensive. Balancing computational costs, accuracy, and specific physics of a required calculation has driven much of excited state method development. Again, while there are hundreds of possible methods to calculate excited state quantum chemistry properties, we will focus here on methods that attempt to capture the effects of post-excitation orbital relaxations as well as on some of the most commonly used excited state methods. Orbital relaxations can be energetically consequential, especially for CT states [29, 47–51], and are also necessary if a low-cost excited state method is to serve the same role as Hartree-Fock theory does in the ground state by providing an orbital-relaxed platform on which to build in correlation effects.

### 1.8.1 Orbital Optimized CIS

Developed by Subotnik et al, Orbital Optimized CIS (OO-CIS) [47] tries to overcome the tendency of CIS to overestimate charge transfer excitation energies by performing a single

orbital relaxation step after the CIS calculation [52]. The method uses a single Newton-Raphson step to optimize orbitals in which the Hessian is approximated as the Hartree-Fock orbital Hessian. This method also introduces $Y$ as a metric to quantify the charge transfer character of a given excitation

$$Y_{ai} = \frac{\partial E_{CIS}}{\partial \theta_{ai}} \tag{1.23}$$

where $\theta$ is a matrix of occupied-virtual orbital rotation parameters and $Y$ is the matrix of CIS energy derivatives with respect to these orbital parameters. Even with a single optimization step, and at twice the computational cost of a regular CIS calculation, this method shows significant correction to the energy overestimation. It puts the performance of CIS for charge transfer states at par with its performance for valence and Rydberg excitations, which previously was 1-2 eV higher. Naturally, this motivates the question whether a fully optimized set of orbitals, can produce more meaningful orbitals, better excitation energies, and whether this full relaxation can be had at a similar mean-field cost.

### 1.8.2 Maximum Overlap Method and Δ-SCF

The Maximum Overlap Method (MOM) modifies the standard SCF procedure by changing the orbitals that are selected. Instead of choosing the Aufbau orbitals, one chooses the orbitals that have the maximum overlap with the previous step [53–62]. In this manner, excited state determinants can be targeted with relative ease with a few user choices. This approach can target high energy states, such as core excitations [56], as well charge transfer and double excitations [57]. In this method, first, an orbital overlap matrix is defined as

$$\boldsymbol{O} = (\boldsymbol{C}^{\text{old}})^{\dagger} \boldsymbol{S} \boldsymbol{C}^{\text{new}} \tag{1.24}$$

then $O_{ij}$ is the overlap between the $i$th old and $j$th new orbitals, and $p_j$ can be defined as the projection of the $j$th new orbital on the old occupied orbital subspace

$$p_j = \sum_{i}^{n} O_{ij} = \sum_{\nu}^{N} \Big[ \sum_{\mu}^{N} \Big( \sum_{i}^{n} \boldsymbol{C}_{i\mu}^{\text{old}} \Big) \boldsymbol{S}_{\mu\nu} \Big] \boldsymbol{C}_{\nu j}^{\text{new}} \tag{1.25}$$

The full set of these projectors can be calculated at barely an additional cost compared to the SCF cycle, and then the $n$ orbitals with the highest $p_j$ are occupied [53]. It provides a very reasonable first guess for excited states, especially the low lying ones [63]. In [53], the authors show that MOM-based approaches to HF, MP2, and B3LYP gave competitive excitation energies with the regular counterparts of CIS, CIS(D), and TDDFT respectively.

Despite the simplicity and elegance of this method, it suffers from two critical issues. First, it relies heavily on the user knowing what orbitals are needed and which orbitals to use. Secondly, in the absence of good virtuals, this method is susceptible to variational collapse to a lower state than desired [64]. There are a number of variations to the basic MOM method, such as IMOM [57], PIMOM [65], etc, which all make modifications to

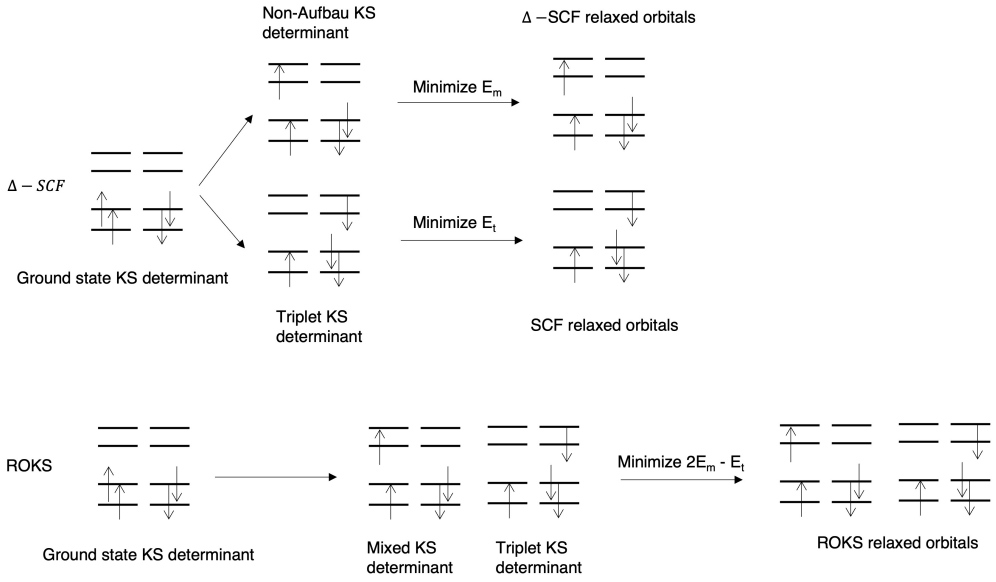


**Figure 1.2:** Graphical description of the Δ-SCF and ROKS orbital relaxation process. The former requires two separate optimizations, followed by a spin purification. ROKS requires only one orbital optimization. Image adapted from [64].

the definition of the projectors based on different properties and in some cases can aid in preventing variational collapse.

The MOM, when it is stable, is an efficient, SCF-style approach to the more general excited state theory known as Δ-SCF. In the Δ-SCF method, first, the ground state calculation is performed. Then, the non-Aufbau configuration is identified and a new SCF cycle is converged in order to fully relax the orbital shapes for this excited configuration [66]. The difference between these two cycles' energy is the excitation energy. However, since only a single determinant is used the excited state wavefunction is not spin-pure, and isn't an eigenstate of the spin operator. If spin purity is desired, a purification post processing step is required. A similar Δ-DFT method can also be constructed, with the DFT SCF cycle and energies used in place of HF [67]. Restricted Open-shell Kohn Sham (ROKS) theory is one method to avoid the secondary spin purification step. In this method, a set of ROKS orbitals are optimized such that the ROKS energy, given below, is stationary.

$$E_S^{ROKS} = 2E_{mix}[(\phi_{ROKS})] - E_T[(\phi_{ROKS})] \tag{1.26}$$

In this equation, $E_{mix}$ is the energy of "mixed" state, which is intermediate between the singlet and triplet. By directly minimizing an expression with the triplet energy fully projected out, the secondary spin purification step is not needed. Fig. 1.2 describes this difference.

In a test set of large organic dyes [67], Δ-DFT showed, on average, $< 1$ eV root mean square deviation (RMSD) for vertical excitation energies for most choice of functionals. The

more spin contaminated the wavefunction was, the more exact exchange was required to achieve an accurate excitation energy. A few years later, the same authors tested the same set of chromophores with ROKS and found ROKS to perform similarly for some functionals, but slightly worse for others [64]. It is important to note that the best performance ($< 0.3$ eV RMSD) was seen for functionals that incorporated long-range corrections (LRC family of functionals), emphasizing the importance of accounting for delocalization in these large chromophores. These works emphasize the importance of functional choice, as well demonstrating that they worked the best for HOMO to LUMO excitations, falling short for higher order excitations.

### 1.8.3 Time Dependent Density Functional Theory

TDDFT, just like DFT for ground states, is the most widely used method for calculating excitation energies and properties [10]. A thorough derivation of TDDFT is presented here [13], and the same authors have examined DFT's shortcomings for charge transfer excitations caused by over-delocalization because of DFT's self interaction error [68, 69]. In brief, TDDFT is a linear response method that starts from the ground state DFT equations and leads to the following eigenvalue equation

$$\begin{bmatrix} A & B \\ B* & A* \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \omega \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \tag{1.27}$$

where $X$ is the excitation amplitude, $Y$ is the deexcitation amplitude, and $A$ and $B$ are described as

$$A = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + (ia|jb) - c_{HF}(ij|ab) + (1 - c_{HF})(ij|f_{xc}|ab) \tag{1.28}$$

$$B = A - (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} \tag{1.29}$$

where $c_{HF}$ is the percentage of exact Hartree-Fock exchange in the chosen potential, and the $f_{xc}$ term is the response of the chosen $xc$ potential.

The relationship between ground state methods, Hartree-Fock and DFT, and their linear response equivalents is shown in Fig.1.3. It's important to note that one can derive CIS with the CI formalism from HF or by setting $B = 0$ in the linear response theory when the density functional is chosen to be 100% HF exchange with no correlation. When the density functional is more general, the CI formalism is no longer valid, but one can still make the $B = 0$ approximation, known as the Tamm-Dancoff Approximation (TDA) [13]. TDA reduces the computational cost by roughly a factor of two, and provides a much more straightforward algebraic path to getting excitation energies [70].

TDDFT provides a relatively robust pathway to calculating valence excitation energies, on average having an error of $0.1 - 0.5$ eV [13]. However, it often gets into trouble in charge transfer excitations, where correct handling of long range exchange is crucial and the spatially separated unpaired electrons are prone to over-delocalization due to DFT's

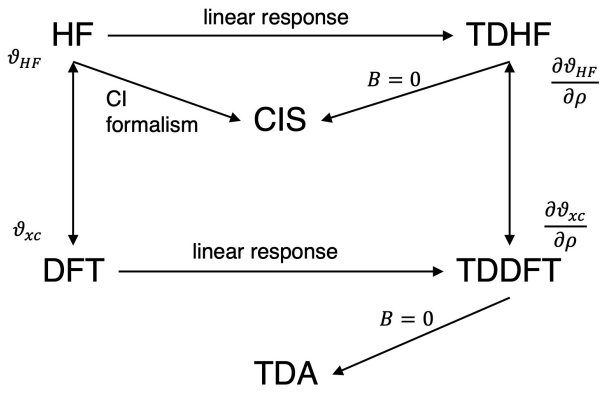


**Figure 1.3:** Relationship between various ground state and response state methods. Figure adapted from [13].

self-interaction error. Despite corrections provided by long-range corrected functionals, the inherent overestimation of charge separation predicted by TDDFT is well established [5, 68, 69, 71, 72].

### 1.8.4 Equation of Motion Coupled Cluster

This method is a response theory formulation to CCSD, and is a $\mathcal{O}(N^6)$ scaling method. Extensively benchmarked [73, 74], this method constructs a linear excitation operator as

$$\hat{\mathcal{R}} = 1 + \sum_i \sum_a r_{ia} a^\dagger i + \frac{1}{4} \sum_{ij} \sum_{ab} r_{ijab} a^\dagger b^\dagger j i \tag{1.30}$$

The corresponding EOM-CCSD wavefunction is described as

$$|\psi_{EOM-CCSD}\rangle = \exp(\hat{T}_1 + \hat{T}_2)(\hat{\mathcal{R}}|\psi\rangle) = \hat{\mathcal{R}}(\exp(\hat{T}_1 + \hat{T}_2)|\psi\rangle) \tag{1.31}$$

The linear excitation operator commutes with the exponential cluster operator, allowing this formulation [75]. It has been noted that the cluster operator is still the operator optimized for the ground state and so some effects are not fully excited-state-specific [2]. For single excitations, approximate orbital-relaxations can be included via the doubles operator within $\hat{R}$, but this only captures such orbital relaxations to linear order. Therefore, the choice of the reference wavefunction is really critical for EOM-CCSD states [76]. As in the ground state, accuracy can be improved through perturbative triples corrections, as in EOM-CCSD(T), although these raise the cost to $\mathcal{O}(N^7)$ [77].

In their landmark benchmarking paper, Thiel and coworkers showed that for single Configuration State Function (CSF) dominated singlet excitation, the maximum mean error was 0.12eV [78]. In similar other large test sets, it was seen that for valence excitations, the

absolute mean error was as low as $0.04 - 0.19$ eV, while for charge transfer excitations it was an average of 0.25 eV, indicating the orbtials being optimized for ground state and not for the excited state starts having a larger impact [79]. In another work [44], the authors showed that, for their test set, to get accuracy down to 0.1eV, EOMEE-CCSD(T) was necessary. Without the perturbative triples, EOM-CCSD overestimated the energies of all types of excitations by $0.1 - 0.3$ eV. In addition to the difficulties with charge transfer excitations, EOM-CC methods present a big problem with regards to efficiency, with only calculations on molecules with 1000 basis functions considered feasible [79].

## 1.9 Outlook

To varying degrees, most of the methods discussed in the previous section retain a reliance on the ground state, either by staying "close" to it through linear response or by employing a ground-state-inspired but spin-broken single-determinant form. One can instead imagine methods that are more fully excited-state-specific and that recreate the ladder of increasingly accurate correlation approaches that all start from a mean-field reference state. In the ground state, HF is on the first rung, and MP2, CC, etc methods lie on higher rungs, with increasing amounts of correlation. As outlined in the next section, this thesis focuses on developing and applying SCF techniques to a possible first rung of an excited-state-specific ladder in which the CIS wave function is improved with full orbital relaxations.

## 1.10 Outline

**Chapter 2**

This chapter introduces a self consistent framework for Excited State Mean Field (ESMF) theory, an excited state-specific analogue to Hartree-Fock theory. Minimally correlated, and based on solving an one-electron mean-field equation self-consistently, ESMF provides excited state energies at roughly twice the cost of a HF calculation. Additionally, just like HF, this excited state method has a commutator criteria for convergence, and is accelerated via DIIS. This work has been published as [80]. Tarini S. Hardikar, Eric Neuscamman; A self-consistent field formulation of excited state mean field theory. J. Chem. Phys. 28 October 2020; 153 (16): 164108. https://doi.org/10.1063/5.0019557

**Chapter 3**

Turning from theory to applications, this chapter focuses on improving our understanding of how well different excited state methods predict the motion of charge in valence and charge transfer excitations when molecules are treated under explicit solvation.

**Chapter 4**

This chapter provides a brief review of this dissertation.

**Chapters 5 and 6**

These chapters contain appendices with additional details in support of Chapters 3 and 4.

# Chapter 2

# A Self Consistent Field Formulation of Excited State Mean Field Theory

## 2.1 Abstract

We show that, as in Hartree Fock theory, the orbitals for excited state mean field theory can be optimized via a self-consistent one-electron equation in which electron-electron repulsion is accounted for through mean field operators. In addition to showing that this excited state ansatz is sufficiently close to a mean field product state to admit a one-electron formulation, this approach brings the orbital optimization speed to within roughly a factor of two of ground state mean field theory. The approach parallels Hartree Fock theory in multiple ways, including the presence of a commutator condition, a one-electron mean-field working equation, and acceleration via direct inversion in the iterative subspace. When combined with a configuration interaction singles Davidson solver for the excitation coefficients, the self consistent field formulation dramatically reduces the cost of the theory compared to previous approaches based on quasi-Newton descent.

## 2.2 Introduction

Hartree Fock (HF) theory [1, 2] is so immensely useful in large part due to the rigorous and convenient link it provides between a qualitatively correct many-electron description and an affordable and more intuitive one-electron equation. The link it makes is rigorous in that, when solved, its one-electron equation guarantees that the many-electron description underneath it is optimal in a variational sense, meaning that the energy is made stationary with respect to changes in the wave function. The link is also convenient, because many-electron properties like the energy can be evaluated in terms of inexpensive one-electron quantities, and because solving a one-electron equation, even one with mean field operators that must be brought to self-consistency, is in most cases easier and less expensive than a direct minimization of the many-electron energy. The fact that this useful link is possible

at all owes much to the simplicity of the Slater determinant many-electron wave function on which HF theory is built. Essentially, the Slater determinant is as close as we can get to a truly mean field, correlation-free Hartree product ansatz while still capturing the important effects of Pauli correlation. Happily, this single step away from a product state does not prevent a useful and intuitive formulation in terms of a self-consistent one-electron equation in which mean field operators account for electron-electron coulomb repulsion.

We will show how excited state mean field (ESMF) theory [49] can also be formulated in terms of a one-electron mean field equation that, when solved self consistently, produces optimal orbitals. As in HF theory, this formulation is possible thanks to the ansatz hewing closely to the mean field limit: ESMF takes only one additional step away from a truly mean field product state by adding the open-shell correlation that arises in an excitation on top of the Pauli correlations already present in the ground state. Perhaps most importantly, the resulting one-electron equation that determines the optimal orbitals can, like the Roothaan equations, be solved by iteratively updating a set of mean field operators until they are self-consistent with the orbital shapes. As we will see, when accelerated by direct inversion in the iterative subspace (DIIS) [25], this self consistent field (SCF) approach brings the orbital optimization cost down to within a factor of two of HF theory, and significantly lowers the overall cost of ESMF theory compared to previous approaches. Given that ESMF offers a powerful platform upon which to construct excited-state-specific correlation theories [33, 34, 81] and that it has recently been shown to out-compete other low-cost methods like configuration interaction singles (CIS) and density functional theory in the prediction of charge density changes [82], this acceleration of the theory and simplification of its implementation should prove broadly useful.

While recent work has provided an improved ability to optimize the ESMF ansatz via the nonlinear minimization of a generalized variational principle (GVP) [33, 82], the current lack of an SCF formulation stands in sharp contrast to the general state of affairs for methods based on Slater determinants. Even in contexts outside of standard HF for ground states, SCF procedures are the norm rather than the exception when it comes to optimizing Slater determinants' orbitals. Indeed, among many others, the $\Delta$SCF [83–86], restricted open-shell Kohn Sham [64, 87], constrained density functional theory [88], ensemble density functional theory [19, 20, 89], projected HF [18], and $\sigma$-SCF [90, 91] methods all favor SCF optimization approaches. Although the direct minimization of a GVP or the norm of the energy gradient [29] offers protection against a Slater determinant's "variational collapse" to the ground state or lower excited states, this rigorous safety comes at some cost to efficiency. It is not for nothing that direct energy minimization methods, although available [16], are not the default HF optimization methods in quantum chemistry codes. In cases where they prove stable, SCF approaches are typically more efficient. In the case of the ESMF anstatz, an SCF approach is also at risk of collapse to an undesired state, but, even in such troublesome cases, a brief relaxation of the orbitals by SCF may still offer a low-cost head start for the direct minimization of a GVP. In cases where an SCF approach to ESMF is stable, history strongly suggests that it will be more efficient than nonlinear minimization. In short, our preliminary data agree with history's suggestion.

## 2.3 Theory

### 2.3.1 Hartree-Fock Theory

To understand how an SCF formulation of ESMF theory comes about, it is useful to first review the formulation of HF theory and in particular how its condition for optimal orbitals can be written as a commutator between a mean field operator and a one-body reduced density matrix (RDM). In HF theory, the energy of the Slater determinant $\Psi_{SD}$ is made stationary with respect to changes in the orbital variables, which is the Slater determinant's approximation of the more general condition that an exact energy eigenstate will have an energy that is stationary with respect to any infinitesimal variation in the wave function. For convenience, and without loss of generality, the molecular orbitals are constrained by Lagrange multipliers to be orthonormal [1].

For Restricted Hartree Fock (RHF), the resulting Lagrangian

$$L_{RHF} = E_{RHF} + 2\,\mathrm{tr}\left[(\boldsymbol{I} - \boldsymbol{C}^T\boldsymbol{S}\boldsymbol{C})\boldsymbol{\epsilon}\right] \tag{2.1}$$

in which $\boldsymbol{C}$ is the matrix whose columns hold the molecular orbital coefficients, $\boldsymbol{S}$ is the atomic orbital overlap matrix, $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{\epsilon}$ is the symmetric matrix of Lagrange multipliers, tr[] is the matrix trace operation, and $E_{RHF}$ is the RHF energy (given below), is then made stationary by setting derivatives with respect to $\boldsymbol{C}$ equal to zero. After some rearrangement [1], this condition can be formulated into the famous Roothaan equations,

$$\big(\boldsymbol{h} + \boldsymbol{W}\left[\boldsymbol{A}\right]\big)\boldsymbol{C} = \boldsymbol{S}\boldsymbol{C}\boldsymbol{\epsilon} \tag{2.2}$$

in which $\boldsymbol{h}$ is the matrix representation of the one-electron components of the Hamiltonian in the atomic orbital basis and $\boldsymbol{W}$ is interpreted as a mean field approximation for electron-electron repulsions. Of course, this mean field repulsion depends on the orbital shapes, causing the operator $\boldsymbol{W}$ to be a function of $\boldsymbol{A}$, the Aufbau determinant's 1-body $\alpha$-spin RDM. In what comes below we will consider RDMs and other matrices in both the atomic orbital (AO) and molecular orbital (MO) bases, and will adopt the notation that a matrix with no superscript (e.g. $\boldsymbol{A}$) refers to the AO representation, while the MO representation is explicitly denoted as such (e.g. $\boldsymbol{A}^{(MO)}$). The closed-shell Aufbau determinant's RDM has the form

$$\boldsymbol{A}^{(MO)} = \boldsymbol{I}_o \qquad \boldsymbol{A} = \boldsymbol{C}\boldsymbol{A}^{(MO)}\boldsymbol{C}^T \tag{2.3}$$

where the matrix $\boldsymbol{I}_o$ has ones on the first $n_o$ elements of its diagonal and zeros elsewhere ($n_o$ is the number of occupied molecular orbitals).

Although in many contexts it is useful to separate the restricted HF (RHF) mean field electron-electron repulsion operator $\boldsymbol{W}[\boldsymbol{A}] = 2\boldsymbol{J}[\boldsymbol{A}] - \boldsymbol{K}[\boldsymbol{A}]$ into its "Coulomb" $\boldsymbol{J}$ and "ex-

change" $\boldsymbol{K}$ components,

$$J[\gamma]_{pq} = \sum_{rs} \gamma_{rs}(rs|pq) \tag{2.4}$$

$$K[\gamma]_{pq} = \sum_{rs} \gamma_{rs}(pr|qs), \tag{2.5}$$

defined here using the two-electron integrals in 1122 order, this separation is not necessary at present and so we will work instead in terms of the combined mean field operator $\boldsymbol{W}$.

Now, while the Roothaan equation has both an intuitive appeal as a one-electron Schrödinger equation and a practical appeal as a convenient setup for an SCF cycle based on the efficient numerical diagonalization of a symmetric generalized eigenvalue problem, it is not the only way to formulate HF theory's central requirement of Lagrangian stationarity. Noting that only the first $n_o$ columns of $\boldsymbol{C}$ affect the ansatz, we can right-multiply Eq. (2.2) by $\boldsymbol{I}_o = \boldsymbol{A}^{(MO)}$ to focus our attention on them while at the same time left-multiplying by $\boldsymbol{C}^T$ to eliminate the overlap matrix, which results in

$$\boldsymbol{F}^{(MO)}\boldsymbol{A}^{(MO)} = \boldsymbol{C}^T\big(\boldsymbol{h}+\boldsymbol{W}\big)\boldsymbol{C}\boldsymbol{A}^{(MO)} = \boldsymbol{\epsilon}\boldsymbol{I}_o \tag{2.6}$$

where we have made the usual definition of the Fock operator.

$$\boldsymbol{F}^{(MO)} = \boldsymbol{C}^T\boldsymbol{F}\boldsymbol{C} = \boldsymbol{C}^T\big(\boldsymbol{h}+\boldsymbol{W}\big)\boldsymbol{C} \tag{2.7}$$

If we ensure that we work in the canonical representation [1], the matrix $\boldsymbol{\epsilon}$ will be diagonal, and so Eq. (2.6) essentially says that the product $\boldsymbol{F}^{(MO)}\boldsymbol{A}^{(MO)}$ must produce a symmetric matrix. We may enforce this requirement by setting the difference between this product and its transpose equal to zero, which leads to a commutator condition for Lagrangian stationarity that can be used as an alternative to the Roothaan equation when optimizing orbitals [25, 92].

$$\big[\,\boldsymbol{C}^T\boldsymbol{F}\boldsymbol{C},\,\boldsymbol{A}^{(MO)}\,\big] = 0 \tag{2.8}$$

If we consider the HF energy expression

$$E_{RHF} = \mathrm{tr}\big[(2\boldsymbol{h}+\boldsymbol{W})\boldsymbol{A}\big] \tag{2.9}$$

alongside the Fock operator definition $\boldsymbol{F} = \boldsymbol{h}+\boldsymbol{W}$, we see a nice connection between the commutator condition and the energy. Specifically, if one halves the one-electron component of the mean field operator whose trace with the density yields the energy, the resulting operator ($\boldsymbol{F}$ in this case) must, when put in the MO basis, commute with the MO basis representation of the density matrix in order for the Lagrangian to be stationary. With this connection pointed out, we now turn our attention to ESMF theory, where a generalization of Eq. (2.8) yields a useful SCF formulation for orbital optimization.

### 2.3.2 Excited State Mean Field Theory

Like HF theory, the energy expression for the ESMF ansatz for a singlet excited state can be written in terms of traces between mean field operators and density-like matrices. In particular, if we take the simple version of the singlet ESMF ansatz in which the Aufbau coefficient is set to zero,

$$|\Psi_{ESMF}\rangle = \sum_{ia} t_{ia} \left|{}^{a\uparrow}_{i\uparrow}\right\rangle + t_{ia} \left|{}^{a\downarrow}_{i\downarrow}\right\rangle, \tag{2.10}$$

where $\boldsymbol{t}$ is the matrix of CIS-like configuration interaction coefficients and $\left|{}^{a\uparrow}_{i\uparrow}\right\rangle$ is the Slater determinant resulting from an $i \to a$ $\alpha$-spin excitation out of the Aufbau determinant (note we do not say the HF determinant, as we are not in the HF MO basis), then the ESMF singlet energy amounts to four traces between mean field operators and density-like matrices.

$$\begin{aligned} E_{ESMF} = \quad & \mathrm{tr}\big[\,(2\boldsymbol{h} + \boldsymbol{W}[\boldsymbol{A}]\,)\,\gamma\,\big] + \mathrm{tr}\big[\,\boldsymbol{W}[\boldsymbol{D}]\,\boldsymbol{A}\,\big] \\ & + \mathrm{tr}\big[\,\boldsymbol{W}[\boldsymbol{T}]\,\boldsymbol{T}^T\,\big] + \mathrm{tr}\big[\,(\boldsymbol{W}[\boldsymbol{T}])^T\,\boldsymbol{T}\,\big] \end{aligned} \tag{2.11}$$

Here $\gamma$ is the one-body alpha-spin RDM for the ESMF ansatz.

$$\gamma^{(MO)} = \boldsymbol{I}_o + \left(\begin{array}{c|c} -\boldsymbol{t}\boldsymbol{t}^T & 0 \\ \hline 0 & \boldsymbol{t}^T\boldsymbol{t} \end{array}\right) \qquad \gamma = \boldsymbol{C}\gamma^{(MO)}\boldsymbol{C}^T \tag{2.12}$$

The matrix $\boldsymbol{A}$ is the Aufbau determinant's one-body RDM, as in Eq. (2.3). The difference between these density matrices we define as $\boldsymbol{D} = \gamma - \boldsymbol{A}$. Finally, $\boldsymbol{T}$ is the non-symmetric matrix that, in its MO representation, has the $\alpha$-spin transition density matrix between the Aufbau determinant and the ESMF ansatz (which is as for CIS just $\boldsymbol{t}$) in its upper-right corner.

$$\boldsymbol{T}^{(MO)} = \left(\begin{array}{c|c} 0 & \boldsymbol{t} \\ \hline 0 & 0 \end{array}\right) \qquad \boldsymbol{T} = \boldsymbol{C}\boldsymbol{T}^{(MO)}\boldsymbol{C}^T \tag{2.13}$$

With the ESMF energy written in terms of one-body mean field operators and density-like matrices, we can now present our central result, in which the stationarity conditions for the ESMF Lagrangian

$$L_{ESMF} = E_{ESMF} + 2\,\mathrm{tr}\left[(\boldsymbol{I} - \boldsymbol{C}^T\boldsymbol{S}\boldsymbol{C})\boldsymbol{\epsilon}\right] \tag{2.14}$$

with respect to orbital variations are written in a one-electron equation that admits an SCF-style solution. We begin, as in HF theory, by setting the (somewhat messy) derivatives $\partial L_{ESMF}/\partial \boldsymbol{C}$ equal to zero. With some care, this condition can be organized into

$$\begin{aligned} & \big(\boldsymbol{h} + \boldsymbol{W}[\boldsymbol{A}]\big)C\gamma^{(MO)} + \boldsymbol{W}[\boldsymbol{D}]CA^{(MO)} \\ & \quad + \boldsymbol{W}[\boldsymbol{T}]C(T^{(MO)})^T + (\boldsymbol{W}[\boldsymbol{T}])^T C T^{(MO)} = \boldsymbol{S}\boldsymbol{C}\boldsymbol{\epsilon} \end{aligned} \tag{2.15}$$

whose structure is similar to but also notably different from the analogous HF expression in Eq. (2.2). The formal difference is that there are now four terms on the left hand side, one for each trace in the energy expression. The practical difference is that the ESMF equation is not an eigenvalue problem, and it is not obvious that it can be reorganized into one due to the incompatible kernels of the matrices $\gamma^{(MO)}$, $A^{(MO)}$, and $T^{(MO)}$.

Thus, it is at present not clear whether this ESMF equation can offer the same spectral information that the Roothaan equation provides for HF. Nonetheless, for orbital optimization, we have found a convenient alternative by transforming this stationary condition into commutator form by following the same steps that took us from Eq. (2.2) to Eq. (2.8) in HF theory. Defining $\boldsymbol{F}_A = \boldsymbol{h} + \boldsymbol{W}[\boldsymbol{A}]$, the result is that the Lagrangian stationary condition can be written as

$$\begin{aligned} 0 = &\left[ \boldsymbol{C}^T \boldsymbol{F}_A \boldsymbol{C},\ \gamma^{(MO)} \right] \\ &+ \left[ \boldsymbol{C}^T \boldsymbol{W}[\boldsymbol{D}] \boldsymbol{C},\ \boldsymbol{A}^{(MO)} \right] \\ &+ \left[ \boldsymbol{C}^T \boldsymbol{W}[\boldsymbol{T}] \boldsymbol{C},\ (\boldsymbol{T}^{(MO)})^T \right] \\ &+ \left[ \boldsymbol{C}^T (\boldsymbol{W}[\boldsymbol{T}])^T \boldsymbol{C},\ \boldsymbol{T}^{(MO)} \right]. \end{aligned} \tag{2.16}$$

It is interesting that the same pattern holds as in the HF case: the commutator condition has one commutator per trace in the energy expression, and the mean field operators (with any one-electron parts halved) are again paired with the same density-like matrices as in the energy traces. We find this pattern especially interesting in light of the fact that it does not simply follow that each trace produces one commutator. Instead, cancellations of terms coming from derivatives on different traces are needed to arrive at the commutators above, and so we do wonder whether this is a happy accident or whether there is an underlying reason to expect such cancellations.

### 2.3.3 Self Consistent Solution

Either way, Eq. (2.16) forms the basis for an efficient SCF optimization of the ESMF orbitals. Assuming that we are a small orbital rotation away from stationarity, we insert the rotation $\boldsymbol{C} \rightarrow \boldsymbol{C}\exp(\boldsymbol{X})$ into our commutator condition and then expand the exponential and drop all terms higher than linear order in the anti-symmetric matrix $\boldsymbol{X}$. The result is a linear equation for $\boldsymbol{X}$ (see Eq. (5.24) in the Appendix) which we solve via the iterative GMRES method. Note that, if desired, one can control the maximum step size in $\boldsymbol{X}$ by simply stopping the GMRES iterations early if the norm of $\boldsymbol{X}$ grows beyond a user-supplied threshold. This may be desirable, as we did after all assume that only a small rotation was needed and our linearization of the equation prevents us from trusting any proposed rotation that is large in magnitude. In parallel to SCF HF theory, which holds $\boldsymbol{F}$ fixed while solving the Roothaan equation for new orbitals, we hold $\boldsymbol{F}_A$, $\boldsymbol{W}[\boldsymbol{D}]$, and $\boldsymbol{W}[\boldsymbol{T}]$ fixed while solving our linear equation. Thus, although the modified GMRES solver is not as efficient as the dense eigenvalue solvers used for HF theory, it remains relatively inexpensive as it does not

does not involve any Fock builds and so does not have to access the two-electron integrals. (*Technical note: in practice, we can speed up the GMRES solver considerably by preconditioning it with a diagonal approximation to the linear transformation that is set to one for* $\boldsymbol{X}$ *elements in the occupied-occupied and virtual-virtual blocks (since these are expected to play little role in the orbital relaxation) and, in the other blocks, replaces* $\boldsymbol{C}^T\boldsymbol{F}_A\boldsymbol{C}$ *with its diagonal, replaces* $\gamma^{(MO)}$ *with* $\boldsymbol{I}_o$*, and neglects* $\boldsymbol{W}[\boldsymbol{D}]$ *and* $\boldsymbol{W}[\boldsymbol{T}]$ *(see Appendix for the explicit form). DIIS is also effective when we take Eq. (2.16) transformed into the AO basis as the error vector and the* $\boldsymbol{F}_A$*,* $\boldsymbol{W}[\boldsymbol{D}]$ *and* $\boldsymbol{W}[\boldsymbol{T}]$ *matrices as the DIIS parameters. We use both of these accelerations in all calculations.*) Only after the linear equation is solved and the orbitals are updated do we rebuild the three mean field operators, and so each overall SCF iteration requires just three Fock builds, which, as they can be done during the same loop over the two-electron integrals, come at a cost that is not much different than HF theory's single Fock build. This arrangement contrasts sharply with the nine Fock builds and two integral loops that are necessary to form the analytic derivative of the energy with respect to $\boldsymbol{C}$ that is used in descent-based orbital optimization [82]. In summary, the ESMF orbitals, like the HF orbitals, can be optimized particularly efficiently via the self-consistent solution of a one-electron mean field equation.

Although this exciting result makes clear that the ESMF ansatz really does hew closely enough to the mean field product-state limit for one-electron mathematics to be of use, there are a number of questions we should now address. First, and we will go into more detail on this point in the next paragraph, is the SCF approach actually faster than descent? The answer, at least in simple systems, is a resounding yes. Second, what of the configuration interaction coefficients $\boldsymbol{t}$? At present, we optimize them in a two-step approach, in which we go back and forth between orbital SCF solutions and CIS calculations (taking care to include the new terms that arise for CIS when not in the HF MO basis) until the energy stops changing. In future, more sophisticated approaches that provide approximate coupling between these optimizations may be possible, as has long been true in multi-reference theory [93]. Third, what physical roles can we ascribe to the different mean field operators that appear in the SCF approach to ESMF? The operator $\boldsymbol{F}_A$ obviously carries the lion's share of the electron-electron repulsion, as it is the only mean field operator derived from a many-electron density matrix. Indeed, $\boldsymbol{W}[\boldsymbol{D}]$ and $\boldsymbol{W}[\boldsymbol{T}]$ represent repulsion from one-electron densities, and so they cannot provide the bulk of the electron-electron repulsion. Thus, we suggest that it is useful to view $\boldsymbol{F}_A$ as a good starting point that includes the various repulsions between electrons not involved in the excitation but that gets the repulsions affected by the excitation wrong. $\boldsymbol{W}[\boldsymbol{D}]$ and $\boldsymbol{W}[\boldsymbol{T}]$ then act as single-electron-density corrections to this starting point. If one considers the simple case in which we ignore all electrons other than the pair involved in the excitation (e.g. consider the HOMO/LUMO excitation in $H_2$), then a close inspection reveals that $\boldsymbol{W}[\boldsymbol{D}]$ eliminates the spurious HOMO-HOMO repulsion that is present in the first trace of the energy expression, while the $\boldsymbol{W}[\boldsymbol{T}]$ terms bring the excited electron pair's repulsion energy into alignment with the actual repulsion energy that results from the singlet's equal superposition of two open-shell determinants.

**Table 2.1:** Convergence of SCF- and GVP-based ESMF for the HOMO/LUMO excitation of cc-pVDZ $H_2O$. Initial values for $\boldsymbol{t}$ and $\boldsymbol{C}$ are set to the two-determinant HOMO/LUMO open shell singlet and the RHF orbitals, respectively. For SCF, the two-step method toggled between CIS and SCF calculations, with CIS going first. As the guess is quite good in this system, the GVP optimization set $\mu = 0$ right away and so amounted to a BFGS minimization of the energy gradient norm. At various points during each optimization (measured both by the cumulative number of loops over the TEIs and by the wall time) we report the energy error $\Delta E$ compared to the fully converged energy. Both calculations used a single core on a 2015 MacBook Air.

| SCF ESMF | | | GVP ESMF | | |
|---|---|---|---|---|---|
| TEI Loops | Time (s) | $\Delta E$ (a.u.) | TEI Loops | Time (s) | $\Delta E$ (a.u.) |
| 10 | 0.007 | 0.062605 | 76 | 0.397 | 0.003761 |
| 20 | 0.025 | 0.000032 | 150 | 0.783 | 0.000654 |
| 30 | 0.033 | 0.000004 | 226 | 1.187 | 0.000184 |
| 40 | 0.054 | 0.000000 | 300 | 1.579 | 0.000001 |

**Table 2.2:** Total time in seconds and number of iterations $n_i$ taken for the orbital optimization in the ground state (for RHF) or the excited state (for SCF-based ESMF) to get within $5\mu E_h$ of its fully converged value. The RHF and ESMF methods rely on the same underlying Fock build code, both use DIIS, and both used one core on a 2015 MacBook Air. For ESMF, only the orbitals are optimized, with $\boldsymbol{t}$ set to the HOMO/LUMO open-shell singlet and the initial guess for $\boldsymbol{C}$ set to the RHF orbitals. For RHF, the eigen-orbitals of the one-electron Hamiltonian were used as the initial guess for $\boldsymbol{C}$. Times do not include the generation of one- and two-electron AO integrals, which are the same for both methods.

| Molecule | Basis | RHF (s) | $n_i$ | ESMF (s) | $n_i$ |
|---|---|---|---|---|---|
| water | cc-pVTZ | 0.087 | 8 | 0.185 | 6 |
| formaldehyde | cc-pVTZ | 0.424 | 11 | 0.862 | 8 |
| ethylene | cc-pVTZ | 0.903 | 8 | 1.735 | 6 |
| toluene | cc-pVDZ | 4.366 | 19 | 6.835 | 11 |

## 2.4 Results

### 2.4.1 Efficiency Comparisons

Returning now to the question of practical efficiency, we report in Table 2.1 the convergence of the energy for the HOMO/LUMO excitation in the water molecule for both SCF-based and GVP descent-based ESMF (note all geometries can be found in the Appendix).

Whether one measures by the number of times the expensive two-electron integral (TEI) access must be performed or by the wall time, the two-step SCF approach is dramatically more efficient than GVP-based descent in this case. (The keen-eyed observer will notice that in the SCF case, the TEI loop count and the wall time do not increase at the same rate, which is due to the CIS iterations having many fewer matrix operations to do as compared to SCF in between each access of the TEIs.) If we focus in on just the orbital optimization, as shown in Table 2.2, we find that the SCF approach for ESMF is almost as efficient as ground state HF theory. In practice, of course, we also want to optimize $\boldsymbol{t}$, and for now we rely on the two-step approach, as used in Table 2.1.

While the SCF approach has clear advantages in simple cases, the GVP is still expected to be essential for cases in which the SCF approach may not be stable. For example, without implementing an interior root solver or freezing an open core (and we have not done either), Davidson-based CIS would be problematic for a core excitation. However, as shown in Table 2.3, a combination of an initial SCF optimization of the orbitals followed by a full GVP optimization of $\boldsymbol{t}$ and $\boldsymbol{C}$ together is quite effective. In this case, the SCF approach brings the energy close to its final value, converging to an energy that is too low by 54 $\mu\mathrm{E}_h$ (remember, excited states do not have any upper bound guarantee, even when a variational principle like energy stationarity or the GVP is in use). From this excellent starting point, the GVP's combined optimization of $\boldsymbol{C}$ and $\boldsymbol{t}$ converges quickly to the final energy, needing just ten gradient evaluations to get within 1 $\mu\mathrm{E}_h$. In contrast, if the initial SCF orbital optimization is omitted, the GVP coupled optimization requires hundreds of gradient evaluations (exactly how many depends on the choice for $\omega$ and how $\mu$ is stepped down to zero) [33] to reach the same level of convergence, and was only able to converge to the correct state at all by setting $\mu$ to 0.5 and $\omega$ 0.08 $\mathrm{E}_h$ lower than the final energy for the initial iterations to avoid converging to a higher-energy core excitation. Especially interesting is the fact that, if we move to the aug-cc-pVTZ basis, the ESMF predictions for the two lowest core excitations in $H_2O$ are 534.3 and 536.2 eV, which are quite close to the experimental values [94] of 534.0 and 535.9 eV and which match the delta between them even more closely. Thus, even in cases where the SCF approach would be difficult to use on its own, it can offer significant benefits in partnership with direct minimization.

### 2.4.2 PYCM

To verify that the benefits of the SCF approach are not confined to smaller molecules, we exhibit its use on a charge transfer state in the PYCM molecule that Subotnik used to demonstrate CIS's bias against charge transfer states [52]. Working in a cc-pVDZ basis for the heavy atoms and 6-31G for hydrogen, we consider the lowest charge transfer state, for which we provide iteration-by-iteration convergence details in the Appendix. In Figure 2.1, we plot the ESMF prediction for the donor and acceptor orbitals, which in this case are just the relaxed HOMO and LUMO orbitals as the $\boldsymbol{t}$ matrix coefficients are strongly dominated by the HOMO→LUMO transition. We see that this state transfers charge from the $\pi$ bonding orbital on the methylated ethylene moiety to the $\pi^*$ orbital on the cyano-

**Table 2.3:** Convergence of the energy for the lowest singlet core excited state of $H_2O$ in the aug-cc-pVDZ basis. Initial values for $\boldsymbol{t}$ and $\boldsymbol{C}$ are set to the two-determinant 1s→LUMO open shell singlet and the RHF orbitals, respectively. An initial SCF optimization converged after 10 iterations (involving one TEI loop each), after which GVP-based BFGS descent (again with $\mu$ set immediately to zero) was started from the SCF result (the GVP requires 2 TEI loops per gradient evaluation). We report the energy error $\Delta E$ compared to the fully converged energy as a function of the cumulative wall time and the cumulative number of TEI loops. The calculation used a single core on a 2015 MacBook Air.

| TEI Loops | Time (s) | $\Delta E$ (a.u.) |
|---|---|---|
| Start with SCF: | | |
| 5 | 0.163 | 0.008698 |
| 10 | 0.267 | -0.000054 |
| Switch to GVP: | | |
| 20 | 0.435 | 0.000002 |
| 30 | 0.604 | 0.000001 |

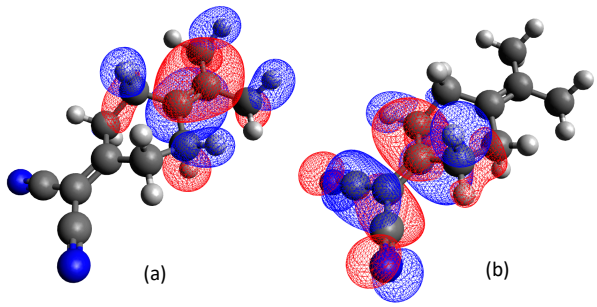


**Figure 2.1:** Donor (a) and acceptor (b) orbitals for the lowest charge transfer state in the PYCM molecule as predicted by ESMF. The excited state SCF calculation took just two and a half times as long as the RHF calculation.

substituted ethylene moiety. Aside from the efficiency of the SCF solver in this case (it takes just two-and-a-half times as long as RHF when using the same Fock build code) it is interesting to compare the prediction against that of CIS, which is the analogous theory when orbital relaxation is ignored. CIS predicts a 7.30 eV excitation energy for the lowest state in which this charge transfer transition plays a significant role, whereas ESMF predicts a 4.82 eV excitation energy. This multiple-eV energy lowering after orbital relaxation serves as a stark reminder of how important these relaxations are for charge transfer states.

## 2.5 Conclusion

In conclusion, orbital optimization in ESMF theory can be formulated in terms of a one-electron equation in which mean field operators provide electron-electron repulsion and which is brought to self-consistency through an efficient iterative process that closely mirrors ground state HF theory. In particular, it is possible to formulate the excited state many-electron energy in terms of four traces between density matrices and mean field operators, and the central commutator condition likewise contains four commutators between these density matrices and their partner mean field operators. In a sense, this is a straightforward extension of the HF case, where only one trace and one commutator are needed. As has long been true for Slater determinants, the SCF approach to the ESMF orbitals appears to be significantly more efficient than quasi-Newton methods, at least in cases where the SCF iteration converges stably to the desired state. Looking forward, it will be interesting to see if, as in the ground state case, the SCF approach admits Kohn-Sham-style density functionals and whether the optimization of the excitation coefficients can be more tightly coupled to the optimization of the orbitals.

### Appendix

See Appendix for additional mathematical details, additional calculation details, and molecular geometries.

### ACKNOWLEDGEMENTS

This work was supported by the Early Career Research Program of the Office of Science, Office of Basic Energy Sciences, the U.S. Department of Energy, grant No. DE-SC0017869. While final timing calculations were carried out on a laptop, many preliminary calculations with earlier pilot code were performed using the Berkeley Research Computing Savio cluster.

*Data Availability Statement* — The data that supports the findings of this study are available within the article and its supplementary material.

# Chapter 3

# Exploring Mulliken Population Changes upon Excitation in Explicitly Solvated Systems

## 3.1 Abstract

The abilities of excited state mean field theory, time dependent density functional theory, and restricted open-shell Kohn-Sham theory to predict the motion of charge in molecular charge transfer states are assessed in the presence of explicit solvent water molecules. Using equation of motion coupled cluster theory as a reference, changes in Mulliken populations between the ground and excited state reveal important differences between these methods. Much more than the other two methods, time dependent density functional theory tends to overestimate the amount of charge that moves. Restricted open-shell Kohn Sham (when used with a range-separated hybrid) and excited state mean field theory are more accurate, and it varies by system which of these two approaches best reproduces the charge motion predicted by equation of motion coupled cluster theory. One noticeable difference between them is that excited state mean field theory avoids density functional theory's self-interaction-induced tendency to artificially delocalize the hole orbital, which we confirm as an issue in one system and see evidence for in another.

## 3.2 Introduction

How the electron charge density distribution within a molecule changes during chemical processes offers significant chemical insight and provides clarity about how the electron cloud participates in and responds to such processes. Many tools have been constructed to analyze charge distributions, and one of the most widely used routes is population analysis [95–101]. Originally most widely used in ground states, these analyses have also been developed for excited state methods [102–107], which allows a user to get, at a glance, a guide to the

degree to which a given excitation transfers charge between different regions of a molecule [108–113].

In ground states, it is well known that the incorporation of correlation effects through methods like density functional theory (DFT) can significantly change orbital shapes and Mulliken populations relative to the predictions of uncorrelated Hartree Fock (HF) theory. [114] This effect is somewhat intuitive, in that electrons that can correlate their motion with each other will incur less electron-electron repulsion as they try to squeeze in closer to the nuclei, and so will manage to do so more effecvtively than in an uncorrelated theory. Beyond Mulliken populations, the difference between DFT orbitals and HF orbitals has also been noted in that employing the former increases accuracy in some correlation treatments like third order Møller-Plesset perturbation theory and some coupled cluster methods. [115, 116]

In the excited state, one would expect a similar situation, with mostly uncorrelated methods like the recently introduced excited state mean field theory [33, 49, 80, 82] (ESMF) presumably producing different Mulliken populations than DFT-based methods such as time-dependent density functional theory [13] (TDDFT) and restricted open-shell Kohn-Sham [67, 87, 117] (ROKS). However, the many differences between these theories make it somewhat difficult to predict how exactly their Mulliken predictions should differ. ESMF and ROKS relax all of the orbitals' shapes following the excitation, but ROKS offers a DFT-based correlation treatment while ESMF, by design, avoids treating most correlation effects. TDDFT, on the other hand, incorporates correlation effects but is a linear response theory that in practice cannot relax all of the orbital shapes post-excitation, [118, 119] although it is at least able to reshape the hole and particle orbitals. Further, TDDFT's well know tendency to overestimate the stability of charge transfer (CT) states [120] suggests that it may be prone to overestimating Mulliken population changes due to spuriously predicting too much CT character. Finally, Mulliken populations in both TDDFT and ROKS could be negatively affected by delocalization error, which is known to be more severe in systems with unpaired electrons [121] and has been observed in ROKS. [82]

Although one can analyze CT states and their Mulliken population changes relative to the ground state in the gas phase, the importance of CT processes in solvated environments makes doing so under the influence of solvent molecules even more interesting. Indeed, it has been widely recognized that gas-phase data is often not sufficient for understanding molecular behavior in solvents. [122–124] As solvation effects can be complex and varied, [123] a variety of quantum chemical approaches have been developed to address them. One category of approach is continuum models, [125–136] which seek to avoid explicitly engaging with the huge numbers of degrees of freedom introduced by even modest numbers solvent molecules (for example, adding 200 octanol solvent molecules adds 16,200 degrees of freedom [131]). However, many first-solvation-shell effects [137] like hydrogen bonding, $\pi$-stacking, and even back-bonding interactions [138] are difficult or impossible to reproduce via continuum models. Indeed, a large body of work has focused on understanding how to accurately simulate these effects [139–145]. One route forward is the inclusion of at least some explicit solvent molecules, [146] which in practice can be performed in tandem with continuum models and/or QM/MM methods [147, 148].

Of course, the downside to the explicit solvent approach is that it can greatly increase the size of the system that one must treat quantum mechanically, although there seems to be little choice in this regard since first-solvation effects are nonnegotiably quantum mechanical. Thus, the challenge of simulating relatively large quantum systems is especially present in this setting. As always, DFT and TDDFT are tempting choices due to their modest computational cost, but extensive benchmarking of TDDFT's performance for organic dyes in isolation and in solvation shows a wide range of errors, [149–154] and improving the solvation description sometimes leads to worse performance compared to experiment. [155] Higher-level methods also face challenges, both because of their cost and sometimes due to difficulties with accuracy. A recent study [155] combining equation of motion (EOM) coupled cluster theory with singles and doubles (EOM-CCSD) with a polarizable continuum model (PCM) showed that EOM-CCSD's error was higher for the same molecule in solvation than in isolation.

With the introduction of ESMF as a low-cost starting point for excited states upon which wholly excited-state-specific correlation methods can be built, [34–36] it is interesting to ask how it performs relative to other low-cost options like TDDFT and ROKS in predicting how charge rearranges during electronic excitations, and in particular CT excitations, in the presence of explicit water molecules. To study this topic, we have implemented a version of ESMF that is connected to the integral engine and Fock build system within Q-Chem, [22] which allows us to evaluate ESMF Mulliken population changes in settings with 20 or more explicit water molecules surrounding a solute. As we will see, many of the expected shortcomings of an uncorrelated method vs correlated methods are present, although we do observe a case in which ESMF successfully avoids a DFT delocalization error. Broadly speaking, we find that ROKS tends to offer more accurate excited state Mulliken populations than ESMF for CT excitations, which are in turn more accurate than those of TDDFT.

## 3.3 Theory

### 3.3.1 Excited State Mean Field Theory

ESMF is a minimally correlated excited-state-specific theory that parallels many aspects of HF theory, including a an effective one-electron working equation involving mean-field operators and self-consistent solutions. [80] ESMF begins with an ansatz similar to configuration interaction singles (CIS) that consists of spin-preserving single excitations off the Aufbau determinant,

$$|\Psi_{ESMF}\rangle = \sum_{ia} t_{ia} \left|{}^{a_\uparrow}_{i_\uparrow}\right\rangle + t_{ia} \left|{}^{a_\downarrow}_{i_\downarrow}\right\rangle, \tag{3.1}$$

where $\boldsymbol{t}$ is the matrix of CIS-like configuration interaction coefficients and $\left|{}^{a\uparrow}_{i\uparrow}\right\rangle$ is the Slater determinant resulting from an $i \rightarrow a$ $\alpha$-spin excitation out of the Aufbau determinant.

Where ESMF departs from CIS is that it variationally relaxes all orbital shapes in the presence of a specific excitation, meaning that its Aufbau determinant differs from the HF Aufbau determinant. Note that, in this study, we explicitly withhold the Aufbau determinant from the ESMF wave function, although in some other studies this has been included.

The energy of the ESMF excited state can be expressed as the sum of four traces between mean-field operators and density-like matrices. [80]

$$E_{ESMF} = \mathrm{tr}\big[(2\boldsymbol{h} + \boldsymbol{W}[\boldsymbol{A}])\gamma\big] + \mathrm{tr}\big[\boldsymbol{W}[\boldsymbol{D}]\boldsymbol{A}\big] + \mathrm{tr}\big[\boldsymbol{W}[\boldsymbol{T}]\boldsymbol{T}^T\big] + \mathrm{tr}\big[(\boldsymbol{W}[\boldsymbol{T}])^T\boldsymbol{T}\big] \tag{3.2}$$

Here $\boldsymbol{W}[\boldsymbol{Z}] = 2\boldsymbol{J}[\boldsymbol{Z}] - \boldsymbol{K}[\boldsymbol{Z}]$ is the "Coulomb minus exchange" mean-field operator built from the one-body alpha-spin reduced density matrix (1-RDM) $\boldsymbol{Z}$, $\gamma$ is the ESMF 1-RDM, and $\boldsymbol{A}$, $\boldsymbol{D} = \gamma - \boldsymbol{A}$, and $\boldsymbol{T}$ are the Aufbau determinant 1-RDM, the difference density, and the non-symmetric alpha-spin transition density matrix, respectively.

$$\gamma^{(MO)} = \boldsymbol{I}_o + \left(\begin{array}{c|c} -\boldsymbol{t}\boldsymbol{t}^T & 0 \\ \hline 0 & \boldsymbol{t}^T\boldsymbol{t} \end{array}\right) \qquad \gamma = \boldsymbol{C}\gamma^{(MO)}\boldsymbol{C}^T \tag{3.3}$$

$$\boldsymbol{T}^{(MO)} = \left(\begin{array}{c|c} 0 & \boldsymbol{t} \\ \hline 0 & 0 \end{array}\right) \qquad \boldsymbol{T} = \boldsymbol{C}\boldsymbol{T}^{(MO)}\boldsymbol{C}^T \tag{3.4}$$

In a full ESMF optimization, the CIS coefficients $\boldsymbol{T}$ and the molecular orbital (MO) coefficients $\boldsymbol{C}$ are optimized simultaneously so that the energy is made stationary with respect to both. One approach to this optimization is a self-consistent field (SCF) method [80] that oscillates between a CIS-like Davidson algorithm for finding $\boldsymbol{T}$ and an SCF orbital optimization. This orbital optimization proceeds by micro-iterations in which the three mean-field operators in Eq. 3.2 are held fixed but $\boldsymbol{C}$ is allowed to vary. These micro-iterations exist within macro-iterations, at the beginning of which the mean-field operators are reevaluated and which can be accelerated by DIIS. An alternative to this SCF approach is the direct nonlinear minimization of of a generalized variational principle (GVP). [33] In the present study, the SCF approach was used for all states except for the back-bonding to $\pi^*$ excitation in formaldehyde and the 4-cyanopyridine states and the , which relied on the GVP approach.

### 3.3.2 Restricted Open-Shell Kohn Sham Method

The ROKS approach seeks to address the spin contamination inherent to single-determinant descriptions of open-shell singlet excited states. To do so, ROKS uses a relationship between the open-shell singlet energy and the energies of the corresponding $S_z = 1$ "high-spin" and $S_z = 0$ "low-spin" single determinants.

$$E^{\mathrm{ROKS}}_{\mathrm{singlet}} = 2\,E_{\mathrm{low-spin}} - E_{\mathrm{high-spin}} \tag{3.5}$$

To achieve excited-state-specific orbital relaxations, ROKS then seeks to make this energy stationary with respect to orbital rotations [22, 64] while using methods like the maximum-overlap method, [53] square gradient minimization, [29, 156, 157] or state targeted energy

projection [48] to help ensure convergence to the desired state. One can also view ROKS as an extension of the vector-coupling approach to restricted open-shell Hartree-Fock (ROHF) theory, as it was originally formulated. [87, 117]

It is worth noting that, in some limits, ROKS and ESMF can be very similar. Specifically, if one chooses the "Hartree Fock" density functional within ROKS (no correlation, 100% exact exchange), then the ROKS ansatz becomes equivalent to an ESMF ansatz in which exactly one CIS coefficient is non-zero. Noting this possibility, we have for each solute molecule performed a transition orbital pair analysis [34] to determine the degree to which this single-CIS-coefficient truncation is reasonable, and have found it to be such a good approximation to the ESMF state in the cases with only a few solvent molecules that we have enforced this condition for all systems, which dramatically accelerates the ESMF optimization. Indeed, using this approach with our Q-Chem-based ESMF implementation of the SCF ESMF optimizer [80] leads to ESMF calculations on PYCM with 20 waters (414 basis functions) that are 10 times faster than Q-Chem-based ROKS calculations that employ a range-separated hybrid functional, with default settings.

### 3.3.3 Time Dependent Density Functional Theory

Given the many resources that construct, explain, and review TDDFT and its strengths and weaknesses [13, 68, 69], we will only briefly highlight two key issues here. First, TDDFT fails to produce the correct $1/r$ asymptotic behavior of the donor/acceptor interaction energy in CT states when using jellium-based exchange functionals and simple hybrids.

This formal issue can be addressed by using range-separated hybrids, which smoothly switch from a traditional hybrid's exchange treatment at short range to 100% HF exchange at long range. Second, an issue facing many simple hybrid functionals is the tendency to underestimate the energies of CT states, often erroneously placing them below the first optically bright state. [120] This issue can be mitigated by including high fractions of HF exchange, although this can impede accuracy for other states. [120]

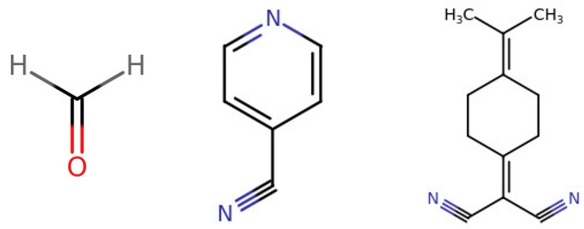


**Figure 3.1:** Lewis structures for formaldehyde (left), 4-cyanopyridine (center), and PYCM (right).

## 3.4 Computational Details

We study the three solute molecules shown in Fig. 3.1. Geometries for these molecules along with small numbers of explicit water molecules were produced using the Conformer-Rotamer Ensemble Sampling Tool (CREST). Specifically, the Quantum Cluster Growth (QCG) algorithm [158, 159] was used to create a solvation geometry, although for cases with too few solvent molecules to create a proper solvation shell it may be more appropriate to call these cluster geometries. These geometries were used in all excited state calculations. CREST, with its QCG algorithm, provides a convenient tool to generate solvated geometries. To produce solvated geometries, the QCG algorithm first applies two wall potentials, one to the solute molecule to keep it centered and a second to the solvent molecules to help prevent clustering and to encourage them to actually surround the solute. With these wall potentials in place, the algorithm alternates between docking and geometry optimization, one solvent molecule at a time, until the user-defined number of solvent molecules have been added. More complete details of this QM and force-field (FF)-based hybrid solvation model are provided by the method's developers [147]. The calculated geometries for all systems studied can be found below in Chapter 6.

Q-Chem was used to perform all EOM-CCSD, CIS, ROKS, and TDDFT calculations [22]. SCF-based ESMF calculations were performed in a development branch of Q-Chem, while GVP-based ESMF calculations were performed with our own standalone code that extracts integrals from PySCF [23]. Three density functionals were tested for both ROKS and TDDFT calculations: B3LYP [160, 161], PBE0 [162], and $\omega$B97X-V [163]. B3LYP and PBE0 were chosen due to their exceedingly wide use in the community. $\omega$B97X-V was chosen as an example of a range-separated hybrid that explicitly accounts for dispersion effects. The 6-31G basis set [164–166] was used in all EOM-CCSD, ESMF, TDDFT, ROKS, and CIS calculations. This modest basis both helps make EOM-CCSD reference calculations affordable and avoids the well-known artifacts that arise for Mulliken populations in large basis sets.

## 3.5 Results

### 3.5.1 Formaldehyde

Formaldehyde was studied both with 8 and with 12 waters. The lowest $n$-to-$\pi^*$ excitation was analyzed in both cases, and an interesting back-bonding-to-$\pi^*$ intermolecular charge transfer state was analyzed in the 12 water case. Beginning with the $n$-to-$\pi^*$ excitation, we found that the relevant lone pair and anti-bonding orbital were the HOMO and LUMO, respectively, in both HF theory and for all three density functionals. This excitation displays some shifting of charge between one end of the formaldehyde and the other, which is to be expected because this excitation moves an electron from the lone pair – which is predominantly localized on the oxygen – into the $\pi^*$ orbital that is shared between the carbon and oxygen but favors

the carbon (which it must do because the corresponding $\pi$ orbital favors the oxygen as electronegativity would predict). For our analysis, we have defined the "donor" region to be formaldehyde's oxygen atom, the "acceptor" to be the $CH_2$ group, and the "water" region to contain all atoms in the water molecules.

As seen in Table 3.1, the reference EOM-CCSD calculations predict that roughly 1/4 and 1/5 of an electron transfer from the donor to acceptor region in the 8 and 12 water cases, respectively. This result reminds us that the presence of solvent really does matter, and provides a basis of comparision for other methods. Regardless of functional, TDDFT shows the anticipated behavior of shifting too much charge, overestimating the degree of CT involved by about a factor of two in both the 8 and 12 water case. It also noticeably overestimates the motion of charge onto the water molecules in the 8 water case, but this is a less severe effect. CIS also overestimates the degree of charge motion, which is behavior that has been seen before in other contexts [47, 52] and is likely due to the fact that it cannot relax the orbital shapes of the other electrons following the excitation. ESMF proves more accurate in its Mulliken population changes than TDDFT or CIS, but ROKS is more accurate still. Both ESMF and ROKS can relax all orbital shapes following the excitation, so the most likely explanation for the difference between them comes from ESMF ignoring most correlation effects. It is also noteworthy that ROKS performs best when using the most sophisticated density functional, although the effect is small for this excitation.

**Table 3.1:** Changes in the Mulliken charges of different regions for the formaldehyde $n \rightarrow \pi^*$ excitation with 8 and 12 waters.

| | ROKS | | | TDDFT | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| # $H_2O$ | B3LYP | PBE0 | $\omega$B97X-V | B3LYP | PBE0 | $\omega$B97X-V | CIS | ESMF | EOM | Region |
| 8 | -0.010 | -0.012 | -0.005 | -0.059 | -0.059 | -0.057 | -0.024 | -0.006 | -0.029 | water |
| | 0.190 | 0.206 | 0.217 | 0.482 | 0.482 | 0.492 | 0.489 | 0.351 | 0.264 | donor |
| | -0.180 | -0.193 | -0.213 | -0.424 | -0.423 | -0.434 | -0.465 | -0.346 | -0.235 | acceptor |
| 12 | 0.018 | 0.012 | 0.008 | 0.027 | 0.018 | 0.001 | 0.003 | 0.005 | 0.000 | water |
| | 0.133 | 0.146 | 0.163 | 0.369 | 0.376 | 0.403 | 0.432 | 0.290 | 0.187 | donor |
| | -0.151 | -0.158 | -0.171 | -0.396 | -0.395 | -0.403 | -0.435 | -0.295 | -0.187 | acceptor |

In the 12 water case, EOM-CCSD predicts that the next-lowest-energy singlet excitation after the $n \rightarrow \pi^*$ excitation is a somewhat curious back-bonding$\rightarrow \pi^*$ intermolecular CT transition. The back-bonding orbital, seen most clearly in the left-hand image of Fig. 3.2, results from a mixing of the out-of-plane lone pair orbital on the water molecule directly below the formaldehyde and the unoccupied $\pi^*$ orbital. Although the $\pi^*$'s shape is somewhat distorted by the interaction, the back-bonding orbital is essentially a constructively interfering linear combination of these orbitals in which the lone pair's upper lobe has same-sign overlap with the $\pi^*$'s lower right lobe. Note that, if you ignore the bottom-most lobe of the back-bonding orbital, the remaining lobes have the characteristic shape of a $\pi^*$ orbital, but one that has been distorted and shifted off of the CO axis by its interaction with the water orbital. As seen in the right-hand image of Fig. 3.2, the excitation in question moves the

electron into an orbital that is more predominately $\pi^*$ in its character, although close inspection of the bottom-right part of this orbital shows that it has a small amount of destructively interfering character between the $\pi^*$ and the water lone pair. So, roughly speaking, these two orbitals, which in this CT excitation are the hole and particle orbitals, are constructive and destructive linear combinations of the water lone pair and the $\pi^*$ orbital.

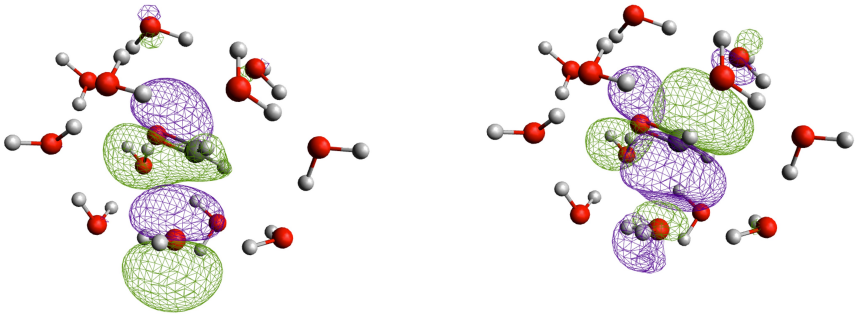

**Figure 3.2:** The hole (at left) and particle (at right) orbitals for the back-bonding to $\pi^*$ excitation in formaldehyde with 12 waters as calculated by ESMF.

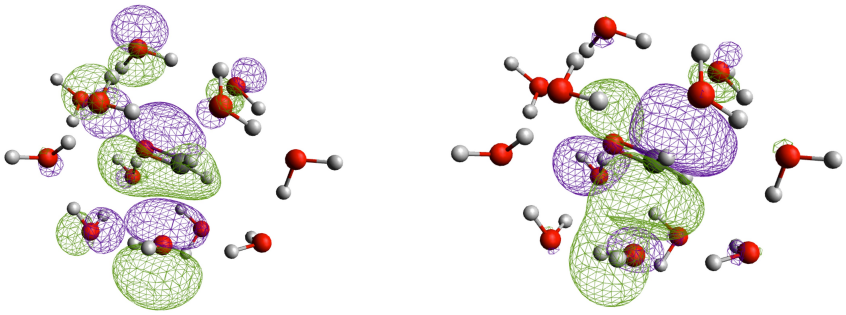

**Figure 3.3:** The hole (at left) and particle (at right) orbitals for the back-bonding to $\pi^*$ excitation in formaldehyde with 12 waters as calculated by ROKS with the $\omega$B97X-V functional.

Turning to Fig. 3.3, we see that ROKS with the $\omega$B97X-V functional makes a similar prediction for the hole and particle orbitals for this transition, except that it delocalizes the hole orbital over additional water molecules. By defining the "donor" region for this transition as the below-the-formaldehyde water molecule that is involved in the back-bonding, the "acceptor" region as the formaldehyde, and the "bystander water" region as all of the other water molecules, we can compare Mulliken changes to those of EOM-CCSD to find out if this hole delocalization is physical or if it is an artificial delocalization driven by DFT's self-interaction error. As seen in Table 3.2, both EOM-CCSD and ESMF predict that the charge on the bystander waters essentially does not change, which is consistent with the fact

that the bystander waters show little to no participation in the hole and particle orbitals in Fig. 3.2. ROKS, on the other hand, predicts that about 1/10 of an electron is removed from the bystander waters during this excitation, which is consistent with its hole orbital being visibly delocalized over multiple waters in Fig. 3.3. Thus, it appears that this delocalization in ROKS is a spurious self-interaction error that ESMF avoids thanks to its exact, wave-function-based treatment of exchange effects.

Overall, no low-level methods do particularly well at predicting the amount of donor-to-acceptor charge transfer in this excitation. ESMF underestimates the charge transfer, while ROKS with the the $\omega$B97X-V functional overestimates it while also showing spurious hole delocalization. Both of these methods dramatically out-perform PBE0 and B3LYP, however, for which we were unable to converge ROKS calculations owing to the very poor initial orbital shapes in the corresponding grounds states. Both PBE0 and B3LYP predict that the back-bonding orbital is strongly mixed with the lone pair orbitals on all the waters, and their TDDFT results show multiple spurious low-lying CT states that involve other waters. One route that will be interesting to explore in future is whether a second order perturbation theory treatment of correlation effects atop an ESMF starting point would produce natural orbitals that more closely match the predictions of EOM-CCSD. This route is promising because it would both provide correlation corrections to the orbital shapes, which is the key missing effect in ESMF, while continuing to avoid the spurious delocalization brought on by DFT's self-interaction error.

**Table 3.2:** Changes in the Mulliken charges of different regions for the formaldehyde back-bonding$\rightarrow \pi^*$ excitation with 12 waters.

| # $H_2O$ | ROKS | ESMF | EOM | Region |
|---|---|---|---|---|
| 12 | 0.096 | -0.010 | -0.012 | bystander waters |
| | 0.388 | 0.161 | 0.341 | donor |
| | -0.484 | -0.151 | -0.329 | acceptor |

### 3.5.2 4-cyanopyridine

For the 4-cyanopyridine charge analysis, we define the "donor" region to be the in-ring nitrogen atom, its two adjacent carbons, and their hydrogens. The remaining four carbons, two hydrogens, and one nitrogen in the solute are defined as the "acceptor" region, while the atoms in water molecules are defined as the "water" region. As seen in Fig. 3.4, the excitation in question is a HOMO-to-LUMO $n \rightarrow \pi^*$ transition. Although TDDFT found this transition to have significant CT character with all three functionals, EOM-CCSD, ESMF, CIS, and ROKS find no significant net transfer of charge, as seen in Table 3.3. Compared to EOM-CCSD, ESMF and ROKS are much more accurate than TDDFT, with CIS more accurate still. This makes the 4-cyanopyridine system interesting in that it is the only system in which CIS does not over-estimate the degree of CT. The reason is likely that this state, despite

what TDDFT says, is not in fact a CT state and so CIS's inability to relax the rest of the orbital shapes has little to no effect on Mulliken populations, unlike in a real CT state like the one we will see in the next section. Thus, 4-cyanopyridine continues the narrative that TDDFT tends to over-estimate the motion of charge, an issue that ROKS and ESMF are less prone to.

**Table 3.3:** Changes in Mulliken charges during the 4-cyanopyridine $n \rightarrow \pi^*$ excitation in gas phase and with 4 explicit waters. Note that ROKS PBE0 did not converge for the latter case.

| | ROKS | | | TDDFT | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| # $H_2O$ | B3LYP | PBE0 | $\omega$B97X-V | B3LYP | PBE0 | $\omega$B97X-V | CIS | ESMF | EOM | Region |
| 0 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | water |
| | 0.033 | 0.045 | 0.045 | 0.404 | 0.375 | 0.246 | 0.032 | 0.093 | -0.004 | donor |
| | -0.061 | -0.076 | -0.074 | -0.155 | -0.140 | -0.068 | -0.065 | -0.065 | -0.040 | acceptor |
| 4 | -0.005 | - | -0.005 | -0.007 | -0.006 | -0.002 | 0.000 | -0.006 | 0.002 | water |
| | 0.097 | - | 0.089 | 0.477 | 0.447 | 0.303 | 0.051 | 0.077 | 0.023 | donor |
| | -0.092 | - | -0.084 | -0.470 | -0.441 | -0.301 | -0.051 | -0.071 | -0.024 | acceptor |

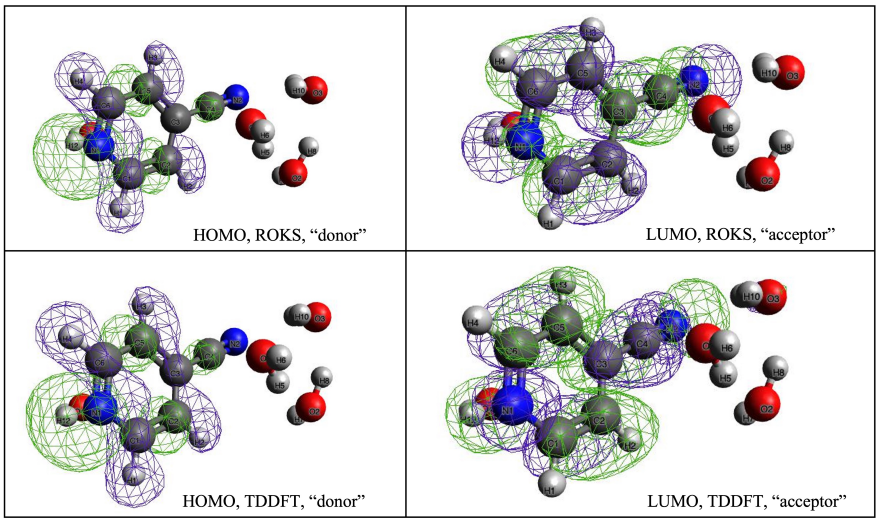


**Figure 3.4:** Participating HOMO and LUMO orbitals for ROKS and TDDFT for the $\omega$B97X-V functional for 4-cyanopyridine with a 4 water solvation shell. Top left: HOMO calculated by ROKS. Top right: LUMO calculated by ROKS. Bottom left: HOMO calculated by TDDFT. Bottom right: LUMO calculated by TDDFT.

### 3.5.3 PYCM

The lowest-lying CT state in PYCM was analyzed in the presence of 4, 10, and 20 water molecules. For illustration, the most relevant HF orbitals are shown for the 10 water case in

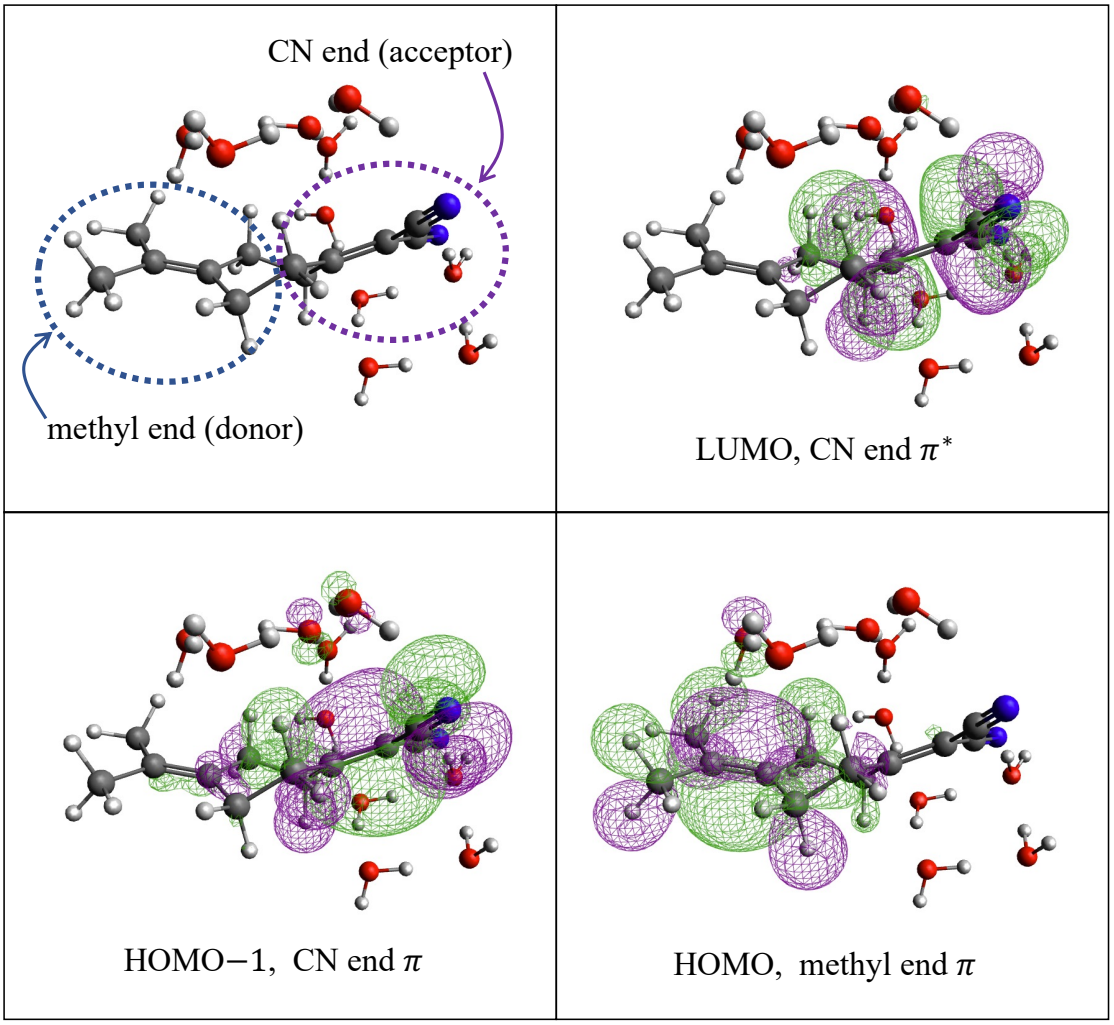


**Figure 3.5:** Relevant HF orbitals in PYCM with 10 water molecules.

**Table 3.4:** Changes in Mulliken charges for the PYCM CT state with 4, 10, and 20 water molecules. Note that ROKS PBE0 did not converge for the largest case. The Ref. data is EOM-CCSD in the 4 water case and CASSCF in the 10 water case.

| | ROKS | | | TDDFT | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| # $H_2O$ | B3LYP | PBE0 | $\omega$B97X-V | B3LYP | PBE0 | $\omega$B97X-V | CIS | ESMF | Ref. | Region |
| 4 | 0.004 | 0.003 | -0.014 | 0.010 | 0.009 | 0.031 | 0.002 | -0.014 | -0.003 | water |
| | 0.122 | 0.127 | 0.502 | 0.905 | 0.902 | 0.605 | 0.725 | 0.536 | 0.578 | donor |
| | -0.126 | -0.130 | -0.488 | -0.914 | -0.911 | -0.636 | -0.727 | -0.522 | -0.575 | acceptor |
| 10 | 0.142 | 0.027 | -0.008 | 0.024 | 0.022 | 0.028 | 0.007 | -0.014 | -0.010 | water |
| | -0.004 | -0.001 | 0.388 | 0.907 | 0.906 | 0.461 | 0.691 | 0.534 | 0.708 | donor |
| | -0.138 | -0.026 | -0.380 | -0.932 | -0.928 | -0.488 | -0.698 | -0.519 | -0.698 | acceptor |
| 20 | -0.020 | - | -0.054 | 0.000 | 0.000 | 0.009 | -0.004 | -0.041 | - | water |
| | 0.171 | - | 0.503 | 0.921 | 0.919 | 0.693 | 0.810 | 0.537 | - | donor |
| | -0.151 | - | -0.449 | -0.922 | -0.919 | -0.701 | -0.806 | -0.496 | - | acceptor |

Fig. 3.5. In the 4 water case, EOM-CCSD, ESMF, and PBE0- and B3LYP-based TDDFT all agree that this state is a simple HOMO to LUMO transition in which the methyl-bearing end's ethyl $\pi$ orbital (the HOMO) acts as the donor and the CN-bearing end's ethyl $\pi^*$ orbital (the LUMO) acts as the acceptor. In the 10 water case, ESMF and the PBE0- and B3LYP-based TDDFT calculations continue to predict the same HOMO-to-LUMO character. EOM-CCSD calculations were not feasible at this system size, but we did compare to a (4e,4o) state-averaged CASSCF calculation in which all 4 of the ethyl $\pi/\pi^*$ orbitals are in the active space, the results of which (see Fig. 3.6) matched the simple HOMO-to-LUMO picture predicted by ESMF. In contrast, when run with the $\omega$B97X-V functional, TDDFT predicts that this CT excitation is a superposition of the HOMO-to-LUMO and HOMO-1-to-LUMO transitions. In other words, it has delocalized the hole orbital across both ends of the molecule. In the 10 water case, ROKS with the $\omega$B97X-V functional shows the same hole orbital delocalization, as seen in Fig. 3.6. Given that ESMF and CASSCF both show clean HOMO-to-LUMO excitations in this 10 water case and that ROKS/$\omega$B97X-V already showed propensity for spurious hole delocalization in formaldehyde, this hole delocalization in 10 water PYCM appears to be another artifact of self-interaction error.

For the Mulliken charge analysis, we define the "donor" region as the six carbon atoms nearest the methyl end of the molecule as well as their hydrogen atoms. The "acceptor" region is defined as the six carbon atoms nearest the CN-bearing end of the molecule along with their hydrogen atoms and the two nitrogen atoms. The "water" region is defined as all oxygen and hydrogen atoms in the water molecules. Using these definitions, Table 3.4 shows this CT excitation's effects on Mulliken charges as predicted by the different methods. ESMF consistently predicts that just over half an electron is transferred between the two ends of the molecule, which closely matches the EOM-CCSD prediction in the 4 water case. ROKS with $\omega$B97X-V predicts less charge transfer, especially in the 10 water case, which is explained by its seemingly spurious hole delocalization discussed above. TDDFT with PBE0 and B3LYP overestimates the degree of charge transfer the same way that it did in formaldehyde, again presumably due to its inability to relax the other orbitals' shapes after the excitation. ROKS

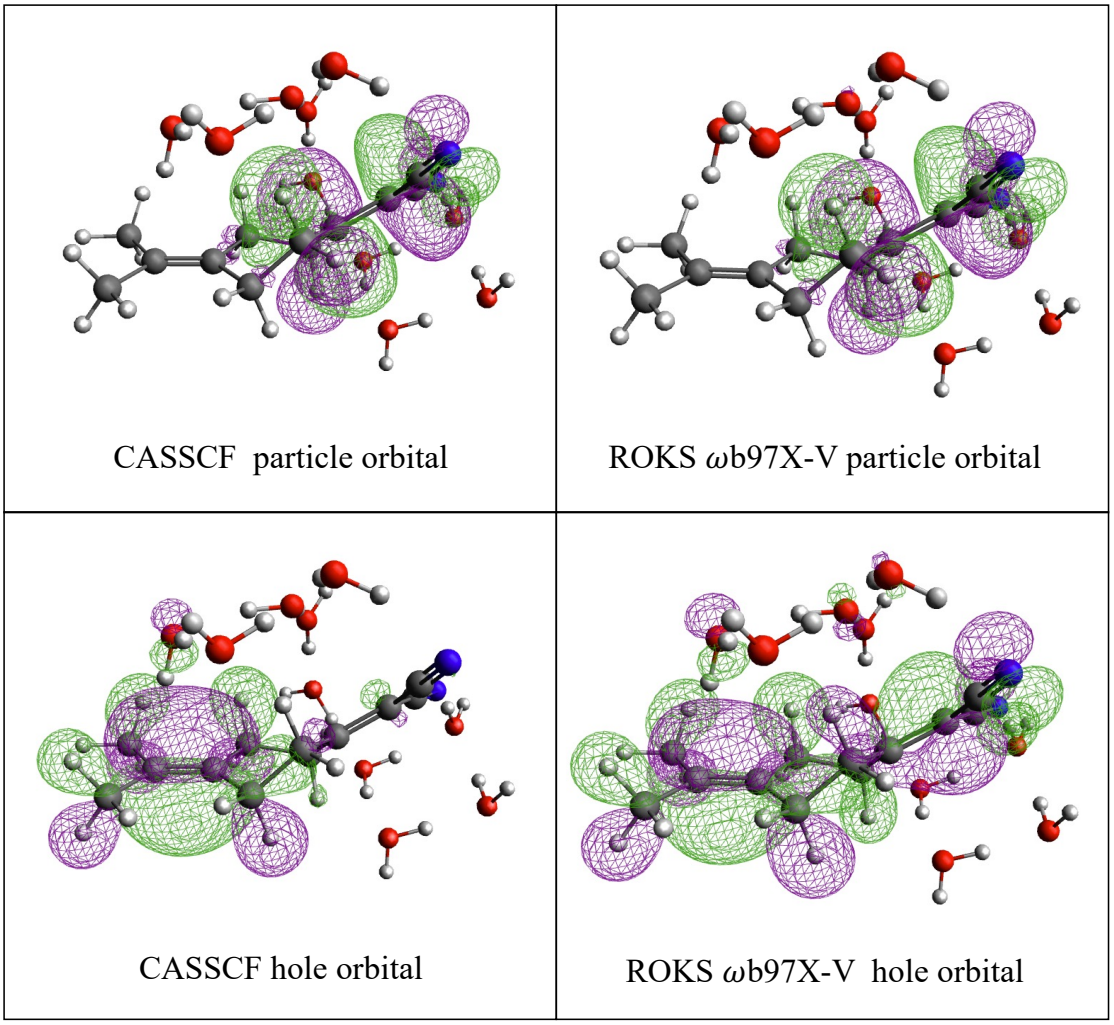


**Figure 3.6:** Comparison of hole and particle orbitals in PYCM with 10 water molecules.

with PBE0 and B3LYP predicts essentially no CT at all, and inspection of the final wave functions reveals that the optimizations collapse onto a spurious non-CT stationary point. This is quite surprising, as the initial guess for these ROKS calculations is the HOMO-to-LUMO transition, and the PBE0 and B3LYP HOMO and LUMO orbitals in all three cases are quite similar to those shown in Fig. 3.5 and so should be an excellent guess for this CT state. Overall, ESMF closely matches EOM-CCSD in the 4 water case and shows the least concerning behavior in the larger cases where we do not have EOM-CCSD data to compare against.

## 3.6 Conclusion

Having analyzed changes in Mulliken populations for excitations in formaldehyde, 4-cyanopyridine, and PYCM with varying numbers of explicit water molecules present, we find clear differences in the behavior of TDDFT, ROKS, and ESMF. Compared to EOM-CCSD reference data, TDDFT, even when using a range-separated hybrid functional, is by far the worst performer with a strong tendency to overestimate the degree of CT. This is true even in the 4-cyanopyridine state that EOM-CCSD reveals to not have any significant CT character at all. ESMF and ROKS, thanks to their ability to perform full orbital relaxations after the excitation, are noticeably more accurate than TDDFT, provided that ROKS employs a range-separated hybrid functional. ROKS results with regular hybrids were significantly less accurate. Range-separated ROKS proved more accurate than ESMF in formaldehyde's $n \rightarrow \pi^*$ state, but showed a clearly spurious delocalization of the hole onto the waters in the back-bonding$\rightarrow \pi^*$ state and evidence of similar delocalization troubles in PYCM, although that case is harder to analyze due to the lack of EOM-CCSD reference data in the 10 and 20 water cases. Thus, it is system dependent whether ESMF or ROKS is preferrable for predicting charge motion, although it is worth noting that range-separated hybrid functionals are significantly more expensive to evaluate and optimize with than ESMF.

Looking forward, it would be desirable to address the fact that ESMF's orbital shapes do not take account of a proper electron correlation treatment. Were this possible, it may offer similar or better accuracies than ROKS across the board while still avoiding spurious hole delocalization. One approach to this goal would be to evaluate natural orbitals for the second order perturbation theory (ESMP2) that has recently been introduced as a post-ESMF correlation correction. That theory has an $N^5$ cost scaling, although the success of local correlation theories for MP2 suggest that this may not remain true indefinitely. At present, it remains difficult to find a mean-field cost excited state method that can consistently reproduce the excited state Mulliken populations of higher level theories.

# Chapter 4

# Conclusion

A self consistent framework for ESMF was developed and introduced as an excited state analogue to Hartree-Fock theory. The formulation provides an one-electron equation with mean-field operators that can be solved self consistently, can be accelerated by DIIS, and has a commutator convergence criteria. At the cost of roughly two Hartree-Fock calculations, the method provides a robust and an economical preliminary calculation to optimize excited state orbitals before introducing higher levels of correlation for higher accuracy. This method serves well for an array of excitations, including valence, core, and charge transfer states. ESMF's performance was also tested for larger solvated molcular systems, and compared to TDDFT, ROKS, CIS, and EOM-CCSD. The method was consistently correct in calculating if and when the solvent water molecules participated in the charge transfer excitation, unlike TDDFT and ROKS. However, for excitations that did not involve the water molecules, ROKS provided a much more accurate quantitative picture of the HOMO to LUMO excitation, indicating that the ESMF, as a minimally correlated theory, does miss crucial physics and overestimates charge transfer. This suggests that higher correlated theories built atop the ESMF framework, such as ESMP2 and ES-CC, provide an excellent pathway to getting correct qualitative and quantitave pictures for charge transfer excitations in solvated systems.

# Bibliography


[1] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola, N.Y.: Dover Publications, 1996.

[2] T. Helgaker, P. Jøgensen, and J. Olsen. *Molecular Electronic Structure Theory*. West Sussex, England: John Wiley and Sons, Ltd, 2000.

[3] Changfeng Fang, Baswanth Oruganti, and Bo Durbeej. "How Method-Dependent Are Calculated Differences between Vertical, Adiabatic, and 0–0 Excitation Energies?" In: *The Journal of Physical Chemistry A* 118.23 (2014). PMID: 24848558, pp. 4157–4171. DOI: 10.1021/jp501974p. URL: https://doi.org/10.1021/jp501974p.

[4] Stephan Kümmel. "Charge-Transfer Excitations: A Challenge for Time-Dependent Density Functional Theory That Has Been Met". In: *Advanced Energy Materials* 7.16 (2017), p. 1700440. DOI: https://doi.org/10.1002/aenm.201700440.

[5] Dávid Mester and Mihály Kállay. "Charge-Transfer Excitations within Density Functional Theory: How Accurate Are the Most Recommended Approaches?" In: *Journal of Chemical Theory and Computation* 18.3 (2022). PMID: 35200021, pp. 1646–1662. DOI: 10.1021/acs.jctc.1c01307.

[6] Patrick Norman and Andreas Dreuw. "Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules". In: *Chemical Reviews* 118.15 (2018). PMID: 29894157, pp. 7208–7248. DOI: 10.1021/acs.chemrev.8b00156.

[7] Mickaël Véril et al. "QUESTDB: A database of highly accurate excitation energies for the electronic structure community". In: *WIREs Computational Molecular Science* 11.5 (2021), e1517. DOI: https://doi.org/10.1002/wcms.1517.

[8] Jiaxin Ning and Donald G. Truhlar. "The valence and Rydberg states of dienes". In: *Phys. Chem. Chem. Phys.* 22 (11 2020), pp. 6176–6183. DOI: 10.1039/C9CP06952F. URL: http://dx.doi.org/10.1039/C9CP06952F.

[9] Donald A. (Donald Allan) McQuarrie. *Quantum chemistry / Donald A. McQuarrie.* eng. 2nd ed. Sausalito, Calif: University Science Books, 2008. ISBN: 9781891389504.

[10] Soumen Ghosh et al. "Combining Wave Function Methods with Density Functional Theory for Excited States". In: *Chemical Reviews* 118.15 (2018). PMID: 30044618, pp. 7249–7292. DOI: 10.1021/acs.chemrev.8b00193.

[11] Narbe Mardirossian and Martin Head-Gordon. "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals". In: *Mol. Phys.* 115.19 (2017), pp. 2315–2372. DOI: 10.1080/00268976.2017.1333644.

[12] Markus Bursch et al. "Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**". In: *Angewandte Chemie International Edition* 61.42 (2022), e202205735. DOI: https://doi.org/10.1002/anie.202205735.

[13] Andreas Dreuw and Martin Head-gordon. "Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules". In: *Sciences-New York* (2005), pp. 4009–4037.

[14] Kieron Burke, Jan Werschnik, and E. K. U. Gross. "Time-dependent density functional theory: Past, present, and future". In: *J. Chem. Phys.* 123 (2005), p. 062206.

[15] Kieron Burke. "Perspective on density functional theory". In: *The Journal of Chemical Physics* 136.15 (Apr. 2012). ISSN: 0021-9606. DOI: 10.1063/1.4704546.

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

[17] Carlos A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and Gustavo E. Scuseria. "Excited electronic states from a variational approach based on symmetry-projected Hartree–Fock configurations". In: *J. Chem. Phys.* 139.22 (2013), p. 224110. DOI: 10.1063/1.4840097.

[18] Carlos A Jiménez-Hoyos et al. "Projected hartree–fock theory". In: *J. Chem. Phys.* 136.16 (2012), p. 164109.

[19] Walter Kohn. "Density-functional theory for excited states in a quasi-local-density approximation". In: *Phys. Rev. A* 34.2 (1986), p. 737.

[20] Andreas K Theophilou. "The energy density functional formalism for excited states". In: *J. Phys. C* 12.24 (1979), p. 5419.

[21] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* 140 (4A 1965), A1133–A1138. DOI: 10.1103/PhysRev.140.A1133.

[22] Evgeny Epifanovsky et al. "Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package". In: *The Journal of Chemical Physics* 155.8 (Aug. 2021). ISSN: 0021-9606. DOI: 10.1063/5.0055522.

[23] Qiming Sun et al. "PySCF: the Python-based simulations of chemistry framework". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 8.1 (2018). ISSN: 17590884.

[24] Peter J. Knowles et al. "Molpro: a general-purpose quantum chemistry program package". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2.2 (2011), pp. 242–253.

[25] Peter Pulay. "Improved SCF convergence acceleration". In: *J. Comput. Chem.* 3.4 (1982), pp. 556–560.

[26] Alejandro J. Garza and Gustavo E. Scuseria. "Comparison of self-consistent field convergence acceleration techniques". In: *The Journal of Chemical Physics* 137.5 (Aug. 2012). ISSN: 0021-9606. DOI: 10.1063/1.4740249.

[27] Roger Fletcher. *Practical Methods of Optimization*. Second. New York, NY, USA: John Wiley & Sons, 1987.

[28] Richard H. Byrd et al. "A Limited Memory Algorithm for Bound Constrained Optimization". In: *SIAM Journal on Scientific Computing* 16.5 (1995), pp. 1190–1208. DOI: 10.1137/0916069.

[29] Diptarka Hait and Martin Head-Gordon. "Excited State Orbital Optimization via Minimizing the Square of the Gradient: General Approach and Application to Singly and Doubly Excited States via Density Functional Theory". In: *Journal of Chemical Theory and Computation* 16.3 (2020). PMID: 32017554, pp. 1699–1710. DOI: 10.1021/acs.jctc.9b01127.

[30] Kevin Carter-Fenk et al. "Electron-Affinity Time-Dependent Density Functional Theory: Formalism and Applications to Core-Excited States". In: *The Journal of Physical Chemistry Letters* 13.41 (2022). PMID: 36215404, pp. 9664–9672. DOI: 10.1021/acs.jpclett.2c02564.

[31] Leonardo A. Cunha et al. "Relativistic Orbital-Optimized Density Functional Theory for Accurate Core-Level Spectroscopy". In: *The Journal of Physical Chemistry Letters* 13.15 (2022). PMID: 35412838, pp. 3438–3449. DOI: 10.1021/acs.jpclett.2c00578.

[32] Eric A. Haugen et al. "Ultrafast X-ray Spectroscopy of Intersystem Crossing in Hexafluoroacetylacetone: Chromophore Photophysics and Spectral Changes in the Face of Electron-Withdrawing Groups". In: *The Journal of Physical Chemistry A* 127.3 (2023). PMID: 36638240, pp. 634–644. DOI: 10.1021/acs.jpca.2c06044.

[33] J. A. R. Shea, E. Gwin, and E. Neuscamman. "A Generalized Variational Principle with Applications to Excited State Mean Field Theory". In: *J. Chem. Theory Comput.* 16 (2020), p. 1526.

[34] Rachel Clune, Jacqueline A. R. Shea, and Eric Neuscamman. "N5-Scaling Excited-State-Specific Perturbation Theory". In: *Journal of Chemical Theory and Computation* 16.10 (2020). PMID: 32816474, pp. 6132–6141. DOI: 10.1021/acs.jctc.0c00308.

[35] Rachel Clune et al. *Studying excited-state-specific perturbation theory on the Thiel set*. 2023. arXiv: 2302.07240 [physics.chem-ph].

[36] Harrison Tuckman and Eric Neuscamman. *An Excited-State-Specific Projected Coupled-Cluster Theory*. 2023. arXiv: 2302.06731 [physics.chem-ph].

[37] Per-Olov Löwdin. "Quantum Theory of Many-Particle Systems. III. Extension of the Hartree-Fock Scheme to Include Degenerate Systems and Correlation Effects". In: *Phys. Rev.* 97 (6 Mar. 1955), pp. 1509–1520. DOI: 10.1103/PhysRev.97.1509. URL: https://link.aps.org/doi/10.1103/PhysRev.97.1509.

[38] Carlos L. Benavides-Riveros, Nektarios N. Lathiotakis, and Miguel A. L. Marques. "Towards a formal definition of static and dynamic electronic correlations". In: *Phys. Chem. Chem. Phys.* 19 (20 2017), pp. 12655–12664. DOI: 10.1039/C7CP01137G. URL: http://dx.doi.org/10.1039/C7CP01137G.

[39] P.J. Knowles and N.C. Handy. "A new determinant-based full configuration interaction method". In: *Chemical Physics Letters* 111.4 (1984), pp. 315–321. ISSN: 0009-2614. DOI: https://doi.org/10.1016/0009-2614(84)85513-X. URL: https://www.sciencedirect.com/science/article/pii/000926148485513X.

[40] Peter J. Knowles and Nicholas C. Handy. "A determinant based full configuration interaction program". In: *Computer Physics Communications* 54.1 (1989), pp. 75–83. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(89)90033-7. URL: https://www.sciencedirect.com/science/article/pii/0010465589900337.

[41] K. A. Brueckner. "Two-Body Forces and Nuclear Saturation. III. Details of the Structure of the Nucleus". In: *Phys. Rev.* 97 (5 Mar. 1955), pp. 1353–1366. DOI: 10.1103/PhysRev.97.1353. URL: https://link.aps.org/doi/10.1103/PhysRev.97.1353.

[42] Jeffrey Goldstone and Nevill Francis Mott. "Derivation of the Brueckner many-body theory". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 239.1217 (1957), pp. 267–279. DOI: 10.1098/rspa.1957.0037.

[43] John F. Stanton and Rodney J. Bartlett. "The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties". In: *The Journal of Chemical Physics* 98.9 (May 1993), pp. 7029–7039. ISSN: 0021-9606. DOI: 10.1063/1.464746.

[44] Péter G. Szalay, Marcel Nooijen, and Rodney J. Bartlett. "Alternative ansätze in single reference coupled-cluster theory. III. A critical analysis of different methods". In: *The Journal of Chemical Physics* 103.1 (1995). Cited by: 95, pp. 281–298. DOI: 10.1063/1.469641. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-36449002268&doi=10.1063%2f1.469641&partnerID=40&md5=7ef5f8cd744d593af3e84e78ce9810c1.

[45] Hideo Sekino and Rodney J. Bartlett. "On the Extensivity Problem in Coupled-Cluster Property Evaluation". In: ed. by Per-Olov Löwdin. Vol. 35. Advances in Quantum Chemistry. Academic Press, 1999, pp. 149–173. DOI: https://doi.org/10.1016/S0065-3276(08)60459-1. URL: https://www.sciencedirect.com/science/article/pii/S0065327608604591.

[46] Hugh G. A. Burton and David J. Wales. "Energy Landscapes for Electronic Structure". In: *Journal of Chemical Theory and Computation* 17.1 (2021). PMID: 33369396, pp. 151–169. DOI: 10.1021/acs.jctc.0c00772.

[47] Xinle Liu et al. "Communication: Adjusting charge transfer state energies for configuration interaction singles: Without any parameterization and with minimal cost". In: *The Journal of Chemical Physics* 136.16 (Apr. 2012). ISSN: 0021-9606. DOI: 10.1063/1.4705757.

[48] Kevin Carter-Fenk and John M. Herbert. "State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions". In: *Journal of Chemical Theory and Computation* 16.8 (2020). PMID: 32644792, pp. 5067–5082. DOI: 10.1021/acs.jctc.0c00502.

[49] Jacqueline A. R. Shea and Eric Neuscamman. "Communication: A mean field platform for excited state quantum chemistry". In: *J. Chem. Phys.* 149.8 (2018), p. 081101.

[50] Emmanuel Giner and Celestino Angeli. "Spin density and orbital optimization in open shell systems: A rational and computationally efficient proposal". In: *The Journal of Chemical Physics* 144.10 (Mar. 2016). 104104. ISSN: 0021-9606. DOI: 10.1063/1.4943187. URL: https://doi.org/10.1063/1.4943187.

[51] Yuan Yao and C. J. Umrigar. "Orbital Optimization in Selected Configuration Interaction Methods". In: *Journal of Chemical Theory and Computation* 17.7 (2021). PMID: 34132530, pp. 4183–4194. DOI: 10.1021/acs.jctc.1c00385.

[52] Joseph E. Subotnik. "Communication: Configuration interaction singles has a large systematic bias against charge-transfer states". In: *J. Chem. Phys.* 135 (2011).

[53] Andrew T B Gilbert, Nicholas A Besley, and Peter M W Gill. "Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method". In: *J. Phys. Chem. A* 112(50) (2008), pp. 13164–13171.

[54] F. Weinhold and T. Brunck. "The principle of maximum overlap". In: *J. Am. Chem. Soc.* 98 (1976), p. 3745.

[55] H. W. Meldner and J. Perez. "Maximum-Ovelap Orbitals". In: *Phys. Rev. C* 7 (1973), p. 2158.

[56] N. A. Besley, A. T. B. Gilbert, and P. M. W. Gill. In: *J. Chem. Phys.* 130 (12 2009), p. 124308.

[57] Giuseppe M. J. Barca, Andrew T. B. Gilbert, and Peter M. W. Gill. "Simple Models for Difficult Electronic Excitations". In: *Journal of Chemical Theory and Computation* 14.3 (2018). PMID: 29444408, pp. 1501–1509. DOI: 10.1021/acs.jctc.7b00994.

[58] F. Liu and C. G. Zhan. "Maximum overlap method and the bond strength". In: *Int. J. Quantum Chem.* 32 (1987), p. 1.

[59] D. H. Kobe. "Maximum-Overlap Orbitals, an Energy Variational Principle, and Perturbation Theory". In: *Phys. Rev. C* 3 (1971), p. 417.

[60] J. Cioslowski and M. Challacombe. "Maximum similarity orbitals for analysis of the electronic excited states". In: *Int. J. Quantum Chem.* 40 (1991), p. 81.

[61] Z. B. Maksić, M. Eckert-Maksić, and M. Randić. "Correlation between CH and CC spin-spin coupling constants and s character of hybrids calculated by the maximum overlap method". In: *Theor. Chim. Acta* 22 (1971), p. 70.

[62] R. J. Bartlett and Y. Öhrn. "How quantitative is the concept of maximum overlap?" In: *Theor. Chim. Acta* 21 (1971), p. 215.

[63] Giuseppe M. J. Barca, Andrew T. B. Gilbert, and Peter M. W. Gill. "Communication: Hartree-Fock description of excited states of H2". In: *The Journal of Chemical Physics* 141.11 (Sept. 2014). ISSN: 0021-9606. DOI: 10.1063/1.4896182.

[64] Tim Kowalczyk et al. "Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach". In: *J. Chem. Phys.* 138 (2013).

[65] Hector H. Corzo et al. "Using projection operators with maximum overlap methods to simplify challenging self-consistent field optimization". In: *Journal of Computational Chemistry* 43.6 (2022), pp. 382–390. DOI: https://doi.org/10.1002/jcc.26797.

[66] T.Ziegler, A. Rauk, and E. J. Baerends. In: *Theor. Chim. Acta* 43 (1977), p. 261.

[67] Tim Kowalczyk, Shane R. Yost, and Troy Van Voorhis. "Assessment of the SCF density functional theory approach for electronic excitations in organic dyes". In: *The Journal of Chemical Physics* 134.5 (Feb. 2011). ISSN: 0021-9606. DOI: 10.1063/1.3530801.

[68] Andreas Dreuw and Martin Head-Gordon. "Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin − Bacteriochlorin and Bacteriochlorophyll − Spheroidene Complexes". In: *Journal of the American Chemical Society* 126.12 (2004). PMID: 15038755, pp. 4007–4016. DOI: 10.1021/ja039556n.

[69] Andreas Dreuw, Jennifer L. Weisman, and Martin Head-Gordon. "Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange". In: *The Journal of Chemical Physics* 119.6 (July 2003), pp. 2943–2946. ISSN: 0021-9606. DOI: 10.1063/1.1590951. URL: https://doi.org/10.1063/1.1590951.

[70] So Hirata and Martin Head-Gordon. "Time-dependent density functional theory within the Tamm–Dancoff approximation". In: *Chemical Physics Letters* 314.3 (1999), pp. 291–299. ISSN: 0009-2614. DOI: https://doi.org/10.1016/S0009-2614(99)01149-5.

[71] Andreas Dreuw, Graham R. Fleming, and Martin Head-Gordon. "Charge-Transfer State as a Possible Signature of a Zeaxanthin−Chlorophyll Dimer in the Non-photochemical Quenching Process in Green Plants". In: *The Journal of Physical Chemistry B* 107.27 (2003), pp. 6500–6503. DOI: 10.1021/jp034562r.

[72] David J. Tozer et al. "Does density functional theory contribute to the understanding of excited states of unsaturated organic compounds?" In: *Molecular Physics* 97.7 (1999), pp. 859–868. DOI: 10.1080/00268979909482888.

[73] Piotr Piecuch, Jared A. Hansen, and Adeayo O. Ajala. "Benchmarking the completely renormalised equation-of-motion coupled-cluster approaches for vertical excitation energies". In: *Molecular Physics* 113.19-20 (2015), pp. 3085–3127. DOI: 10.1080/00268976.2015.1076901.

[74] J. Sous, P. Goel, and M. Nooijen. "Similarity transformed equation of motion coupled cluster theory revisited: a benchmark study of valence excited states". In: *Molecular Physics* 112.5-6 (2014), pp. 616–638. DOI: 10.1080/00268976.2013.847216.

[75] Anna I Krylov. "Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: the Hitchhiker's guide to Fock space." In: *Annu. Rev. Phys. Chem.* 59 (2008), pp. 433–462. ISSN: 0066-426X. DOI: 10.1146/annurev.physchem.59.032607.093602.

[76] Jesse J. Lutz et al. "Reference dependence of the two-determinant coupled-cluster method for triplet and open-shell singlet states of biradical molecules". In: *The Journal of Chemical Physics* 148.16 (Apr. 2018). ISSN: 0021-9606. DOI: 10.1063/1.5025170. URL: https://doi.org/10.1063/1.5025170.

[77] László Gyevi-Nagy, Mihály Kállay, and Péter R. Nagy. "Accurate Reduced-Cost CCSD(T) Energies: Parallel Implementation, Benchmarks, and Large-Scale Applications". In: *Journal of Chemical Theory and Computation* 17.2 (2021). PMID: 33400527, pp. 860–878. DOI: 10.1021/acs.jctc.0c01077.

[78] Marko Schreiber et al. "Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3". In: *The Journal of Chemical Physics* 128.13 (Apr. 2008). 134110. ISSN: 0021-9606. DOI: 10.1063/1.2889385. URL: https://doi.org/10.1063/1.2889385.

[79] Róbert Izsák. "Single-reference coupled cluster methods for computing excitation energies in large molecules: The efficiency and accuracy of approximations". In: *WIREs Computational Molecular Science* 10.3 (2020), e1445. DOI: https://doi.org/10.1002/wcms.1445.

[80] Tarini S. Hardikar and Eric Neuscamman. "A self-consistent field formulation of excited state mean field theory". In: *The Journal of Chemical Physics* 153.16 (Oct. 2020). ISSN: 0021-9606. DOI: 10.1063/5.0019557.

[81] L. Zhao and E. Neuscamman. "Density Functional Extension to Excited-State Mean-Field Theory". In: *J. Chem. Theory Comput.* 16 (2020), p. 164.

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

[83] Paul S Bagus. "Self-consistent-field wave functions for hole states of some Ne-like and Ar-like ions". In: *Phys. Rev.* 139.3A (1965), A619.

[84] Hsiang-lin Hsu, Ernest R Davidson, and Russell M Pitzer. "An SCF method for hole states". In: *J. Chem. Phys.* 65.2 (1976), pp. 609–613.

[85] A Naves de Brito et al. "A theoretical study of x-ray photoelectron spectra of model molecules for polymethylmethacrylate". In: *J. Chem. Phys.* 95.4 (1991), pp. 2965–2974.

[86] Nicholas A Besley, Andrew TB Gilbert, and Peter MW Gill. "Self-consistent-field calculations of core excited states". In: *J. Chem. Phys.* 130.12 (2009), p. 124308.

[87] M. Filatov and S. Shaik. "A spin-restricted ensemble-referenced Kohn-Sham method and its application to diradicaloid situations". In: *Chem. Phys. Lett.* 304 (1999), pp. 429–437.

[88] Qin Wu and Troy Van Voorhis. "Direct optimization method to study constrained systems within density-functional theory". In: *Phys. Rev. A* 72.2 (2005), p. 024502.

[89] Eberhardt KU Gross, Luiz N Oliveira, and Walter Kohn. "Density-functional theory for ensembles of fractionally occupied states. I. Basic formalism". In: *Phys. Rev. A* 37.8 (1988), p. 2809.

[90] Hong-Zhou Ye et al. "$\sigma$-SCF: A direct energy-targeting method to mean-field excited states". In: *J. Chem. Phys* 147.21 (2017), p. 214104.

[91] H. Ye and T. Van Voorhis. "Half-Projected $\sigma$ Self-Consistent Field For Electronic Excited States". In: *J. Chem. Theory Comput.* 15(5) (2019), pp. 2954–2964.

[92] R. McWeeny. *Methods of Molecular Quantum Mechanics*. London: Academic Press, 1996.

[93] David A Kreplin, Peter J Knowles, and Hans-Joachim Werner. "MCSCF optimization revisited. II. Combined first-and second-order orbital optimization for large molecules". In: *J. Chem. Phys.* 152.7 (2020), p. 074102.

[94] J Schirmer et al. "K-shell excitation of the water, ammonia, and methane molecules using high-resolution photoabsorption spectroscopy". In: *Phys. Rev. A* 47.2 (1993), p. 1136.

[95] Per-Olov Löwdin. "On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals". In: *The Journal of Chemical Physics* 18.3 (Dec. 2004), pp. 365–375. ISSN: 0021-9606. DOI: 10.1063/1.1747632.

[96] R. S. Mulliken. "Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I". In: *The Journal of Chemical Physics* 23.10 (Dec. 2004), pp. 1833–1840. ISSN: 0021-9606. DOI: 10.1063/1.1740588.

[97] R. S. Mulliken. "Electronic Population Analysis on LCAO-MO Molecular Wave Functions. III. Effects of Hybridization on Overlap and Gross AO Populations". In: *The Journal of Chemical Physics* 23.12 (Dec. 2004), pp. 2338–2342. DOI: 10.1063/1.1741876.

[98] R. S. Mulliken. "Electronic Population Analysis on LCAO-MO Molecular Wave Functions. IV. Bonding and Antibonding in LCAO and Valence-Bond Theories". In: *The Journal of Chemical Physics* 23.12 (Dec. 2004), pp. 2343–2346. ISSN: 0021-9606. DOI: 10.1063/1.1741877.

[99] R. S. Mulliken. "Criteria for the Construction of Good Self-Consistent-Field Molecular Orbital Wave Functions, and the Significance of LCAO-MO Population Analysis". In: *The Journal of Chemical Physics* 36.12 (July 2004), pp. 3428–3439. ISSN: 0021-9606. DOI: 10.1063/1.1732476.

[100] R. S. Mulliken. "Electronic Population Analysis on LCAO–MO Molecular Wave Functions. II. Overlap Populations, Bond Orders, and Covalent Bond Energies". In: *The Journal of Chemical Physics* 23.10 (Dec. 2004), pp. 1841–1846. ISSN: 0021-9606. DOI: 10.1063/1.1740589.

[101] F. L. Hirshfeld. "Bonded-atom fragments for describing molecular charge densities". In: *Theoretica chimica acta* 44 (1977), pp. 129–138.

[102] Jie Liu and WanZhen Liang. "Analytical approach for the excited-state Hessian in time-dependent density functional theory: Formalism, implementation, and performance". In: *The Journal of Chemical Physics* 135.18 (Nov. 2011). ISSN: 0021-9606. DOI: 10.1063/1.3659312.

[103] Jie Liu and WanZhen Liang. "Analytical second derivatives of excited-state energy within the time-dependent density functional theory coupled with a conductor-like polarizable continuum model". In: *The Journal of Chemical Physics* 138.2 (Jan. 2013). ISSN: 0021-9606. DOI: 10.1063/1.4773397.

[104] Pavel Pokhilko et al. "Spin-Forbidden Channels in Reactions of Unsaturated Hydrocarbons with O(3P)". In: *The Journal of Physical Chemistry A* 123.2 (2019), pp. 482–491. DOI: 10.1021/acs.jpca.8b10225.

[105] Sergey V. Levchenko, Tao Wang, and Anna I. Krylov. "Analytic gradients for the spin-conserving and spin-flipping equation-of-motion coupled-cluster models with single and double substitutions". In: *The Journal of Chemical Physics* 122.22 (June 2005). ISSN: 0021-9606. DOI: 10.1063/1.1877072.

[106] Johannes Neugebauer and Bernd A. Hess. "Fundamental vibrational frequencies of small polyatomic molecules from density-functional calculations and vibrational perturbation theory". In: *The Journal of Chemical Physics* 118.16 (Apr. 2003), pp. 7215–7225. ISSN: 0021-9606. DOI: 10.1063/1.1561045.

[107] Ester Livshits and Roi Baer. "A well-tempered density functional theory of electrons in molecules". In: *Phys. Chem. Chem. Phys.* 9 (23 2007), pp. 2932–2941. DOI: 10.1039/B617919C.

[108] Rustam Z. Khaliullin, Alexis T. Bell, and Martin Head-Gordon. "Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals". In: *The Journal of Chemical Physics* 128.18 (May 2008). ISSN: 0021-9606. DOI: 10.1063/1.2912041.

[109] G. Del Re, P. Otto, and J. Ladik. "Studies on Charge-Transfer Theory Mulliken Population and Measure of Charge Transfer". In: *Israel Journal of Chemistry* 19.1-4 (1980), pp. 265–271. DOI: https://doi.org/10.1002/ijch.198000030.

[110] Borys Szefczyk, W. Andrzej Sokalski, and Jerzy Leszczynski. "Optimal methods for calculation of the amount of intermolecular electron transfer". In: *The Journal of Chemical Physics* 117.15 (Sept. 2002), pp. 6952–6958. ISSN: 0021-9606. DOI: 10.1063/1.1508367.

[111] Suehiro Iwata. "Absolutely Local Occupied and Excited Molecular Orbitals in the Third-Order Single Excitation Perturbation Theory for Molecular Interaction". In: *The Journal of Physical Chemistry A* 114.33 (2010). PMID: 20429565, pp. 8697–8704. DOI: 10.1021/jp101483t.

[112] Zheng Pei et al. "Elucidating the Electronic Structure of a Delayed Fluorescence Emitter via Orbital Interactions, Excitation Energy Components, Charge-Transfer Numbers, and Vibrational Reorganization Energies". In: *The Journal of Physical Chemistry Letters* 12.11 (2021). PMID: 33705139, pp. 2712–2720. DOI: 10.1021/acs.jpclett.1c00094.

[113] Janice A. Steckel. "Ab Initio Calculations of the Interaction between CO2 and the Acetate Ion". In: *The Journal of Physical Chemistry A* 116.47 (2012). PMID: 23102147, pp. 11643–11650. DOI: 10.1021/jp306446d.

[114] Tapas Kar, János G Ángyán, and AB Sannigrahi. "Comparison of ab initio Hartree-Fock and Kohn- Sham orbitals in the calculation of atomic charge, bond index, and valence". In: *The Journal of Physical Chemistry A* 104.44 (2000), pp. 9953–9963.

[115] Adam Rettig et al. "Third-order Møller–Plesset theory made more useful? The role of density functional theory orbitals". In: *Journal of chemical theory and computation* 16.12 (2020), pp. 7473–7489.

[116] Luke W Bertels, Joonho Lee, and Martin Head-Gordon. "Polishing the gold standard: The role of orbital choice in CCSD (T) vibrational frequency prediction". In: *Journal of chemical theory and computation* 17.2 (2021), pp. 742–755.

[117] Michael Filatov and Sason Shaik. "Spin-restricted density functional approach to the open-shell problem". In: *Chemical Physics Letters* 288.5 (1998), pp. 689–697. ISSN: 0009-2614. DOI: https://doi.org/10.1016/S0009-2614(98)00364-9. URL: https://www.sciencedirect.com/science/article/pii/S0009261498003649.

[118] Tom Ziegler et al. "On the relation between time-dependent and variational density functional theory approaches for the determination of excitation energies and transition moments". In: *J. Chem. Phys.* 130 (2009), p. 154102.

[119] Young Choon Park, Mykhaylo Krykunov, and Tom Ziegler. "On the relation between adiabatic time dependent density functional theory (TDDFT) and the $\Delta$ SCF-DFT method. Introducing a numerically stable $\Delta$ SCF-DFT scheme for local functionals based on constricted variational DFT". In: *Mol. Phys.* 113.13-14 (2015), pp. 1636–1647.

[120] RJ Magyar and S Tretiak. "Dependence of spurious charge-transfer excited states on orbital exchange in TDDFT: large molecules and clusters". In: *Journal of chemical theory and computation* 3.3 (2007), pp. 976–987.

[121] Marcus Lundberg and Per EM Siegbahn. "Quantifying the effects of the self-interaction error in DFT: When do the delocalized states appear?" In: *The Journal of chemical physics* 122.22 (2005), p. 224103.

[122] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. "Quantum Mechanical Continuum Solvation Models". In: *Chemical Reviews* 105.8 (2005). PMID: 16092826, pp. 2999–3094. DOI: 10.1021/cr9904009.

[123] Jacopo Tomasi and Maurizio Persico. "Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent". In: *Chemical Reviews* 94.7 (1994), pp. 2027–2094. DOI: 10.1021/cr00031a013.

[124] Modesto Orozco and F. Javier Luque. "Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems". In: *Chemical Reviews* 100.11 (2000). PMID: 11749344, pp. 4187–4226. DOI: 10.1021/cr990052a.

[125] M. Born. "Volumen und Hydratationswärme der Ionen". In: *Z. Physik 1* (1920). PMID: 11849023. DOI: https://doi.org/10.1007/BF01881023.

[126] C. Amovilli et al. In: *Adv. Quantum Chem.* 32 (1998), p. 227.

[127] B. Mennucci, E. Cancès, and J. Tomasi. In: *J. Phys. Chem. B* 101 (1997), p. 10506.

[128] Benedetta Mennucci. "Polarizable continuum model". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.3 (2012), pp. 386–404.

[129] Lars Onsager. "Electric Moments of Molecules in Liquids". In: *Journal of the American Chemical Society* 58.8 (1936), pp. 1486–1493. DOI: 10.1021/ja01299a050.

[130] John G. Kirkwood. "The Dielectric Polarization of Polar Liquids". In: *The Journal of Chemical Physics* 7.10 (Dec. 2004), pp. 911–919. ISSN: 0021-9606. DOI: 10.1063/1.1750343.

[131] Christopher J. Cramer and Donald G. Truhlar. "Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics". In: *Chemical Reviews* 99.8 (1999). PMID: 11849023, pp. 2161–2200. DOI: 10.1021/cr960149m.

[132] R. Cammi and J. Tomasi. In: *J. Comput. Chem.* 16 (1995), p. 1449.

[133] M. Cossi et al. In: *Chem. Phys. Lett.* 255 (1996), p. 327.

[134] V. Barone and M. Cossi. In: *J. Phys. Chem. A* 102 (1998), p. 1995.

[135] E. Cancès, B. Mennucci, and J. Tomasi. In: *J. Chem. Phys.* 107 (1997), p. 3032.

[136] S. Miertus, E. Scrocco, and J. Tomasi. In: *Chem. Phys.* 55 (1981), p. 117.

[137] *O*. Vol. 3. Chichester ; New York : Wiley, 1982. ISBN: 0471971545.

[138] A Alian et al. "Adsorption of radioisotopes from their organic solutions on filter papers the Sb (III)-nonpolar solvents-system". In: *Journal of radioanalytical and nuclear chemistry* 209.1 (1996), pp. 135–145.

[139] Damien Laage, Thomas Elsaesser, and James T. Hynes. "Water Dynamics in the Hydration Shells of Biomolecules". In: *Chemical Reviews* 117.16 (2017). PMID: 28248491, pp. 10694–10725. DOI: 10.1021/acs.chemrev.6b00765.

[140] H. R. Drew and R. E. Dickerson. "Structure of a B-DNA Dodecamer. III. Geometry of Hydration". In: *J. Mol. Biol.* 151 (1981), p. 535.

[141] D. Vlieghe, J. P. Turkenburg, and L. van Meervelt. "B-DNA at Atomic Resolution Reveals Extended Hydration Patterns". In: *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 55 (1999), p. 1495.

[142] M. C. Wiener and S. H. White. "Structure of a Fluid Dioleoylphosphatidylcholine Bilayer Determined by Joint Refinement of X-Ray and Neutron Diffraction Data. 3. Complete Structure". In: *Biophys. J.* 61 (1992), p. 434.

[143] G. Otting, E. Liepinsh, and K. Wüthrich. "Protein Hydration in Aqueous Solution". In: *Science* 254 (1991), p. 974.

[144] B. Schneider, K. Patel, and H. M. Berman. "Hydration of the Phosphate Group in Double-Helical DNA". In: *Biophys. J.* 75 (1998), p. 2422.

[145] H. E. Alper, D. Bassolinoklimas, and T. R. Stouch. "The Limiting Behavior of Water Hydrating a Phospholipid Monolayer - a Computer Simulation Study". In: *J. Chem. Phys.* 99 (1993), p. 5547.

[146] Ronald M. Levy and Emilio Gallicchio. "COMPUTER SIMULATIONS WITH EXPLICIT SOLVENT: Recent Progress in the Thermodynamic Decomposition of Free Energies and in Modeling Electrostatic Effects". In: *Annual Review of Physical Chemistry* 49.1 (1998). PMID: 9933909, pp. 531–567. DOI: 10.1146/annurev.physchem.49.1.531.

[147] Sebastian Spicher et al. "Automated Molecular Cluster Growing for Explicit Solvation by Efficient Force Field and Tight Binding Methods". In: *Journal of Chemical Theory and Computation* 18.5 (2022). PMID: 35482317, pp. 3174–3189. DOI: 10.1021/acs.jctc.2c00239.

[148] Jin Zhang et al. "Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents". In: *Journal of Chemical Theory and Computation* 13.3 (2017). PMID: 28245118, pp. 1034–1043. DOI: 10.1021/acs.jctc.7b00169.

[149] D. Jacquemin et al. "Extensive TD-DFT benchmark: singlet-excited states of organic molecules". In: *J. Chem. Theory Comput.* 5 (2009), p. 2420.

[150] D. Jacquemin et al. "Assessment of long-range corrected functionals performance for n $\rightarrow$ * transitions in organic dyes". In: *J. Chem. Phys.* 127 (2007), p. 094102.

[151] D. Jacquemin et al. "TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids". In: *J. Chem. Theory Comput.* 4 (2008), p. 123.

[152] D. Jacquemin et al. "TD-DFT assessment of functionals for optical 0–0 transitions in solvated dyes". In: *J. Chem. Theory Comput.* 8 (2012), p. 2359.

[153] C. A. Guido et al. "Electronic excitations in solution: The interplay between state specific approaches and a time-dependent density functional theory description". In: *J. Chem. Theory Comput.* 11 (2015), p. 5782.

[154] Christine M Isborn et al. "The charge transfer problem in density functional theory calculations of aqueously solvated molecules". In: *The Journal of Physical Chemistry B* 117.40 (2013), pp. 12189–12201.

[155] Sijin Ren, Joseph Harms, and Marco Caricato. "An EOM-CCSD-PCM Benchmark for Electronic Excitation Energies of Solvated Molecules". In: *Journal of Chemical Theory and Computation* 13.1 (2017). PMID: 27973775, pp. 117–124. DOI: 10.1021/acs.jctc.6b01053.

[156] Diptarka Hait and Martin Head-Gordon. "Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory". In: *arXiv:1911.04709* (2019).

[157] Diptarka Hait and Martin Head-Gordon. "Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining Sub-electronvolt Error from a Restricted Open-Shell Kohn–Sham Approach". In: *The Journal of Physical Chemistry Letters* 11.3 (2020). PMID: 31917579, pp. 775–786. DOI: 10.1021/acs.jpclett.9b03661.

[158] Christoph Bannwarth et al. "Extended tight-binding quantum chemistry methods". In: *WIREs Comput. Mol. Sci.* 11 (2020), e01493. DOI: 10.1002/wcms.1493. URL: https://dx.doi.org/10.1002/wcms.1493.

[159] Stefan Grimme, Christoph Bannwarth, and Philip Shushkov. "A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (Z=1–86)". In: *J. Chem. Theory Comput.* 13.5 (2017), pp. 1989–2009. DOI: 10.1021/acs.jctc.7b00118. URL: https://dx.doi.org/10.1021/acs.jctc.7b00118.

[160] Axel D. Becke. "Density-functional thermochemistry. III. The role of exact exchange". In: *J. Chem. Phys.* 98.7 (1993), pp. 5648–5652. ISSN: 00219606.

[161] P. J. Stephens et al. "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields". In: *The Journal of Physical Chemistry* 98.45 (1994), pp. 11623–11627. DOI: 10.1021/j100096a001.

[162] Carlo Adamo and Vincenzo Barone. "Toward reliable density functional methods without adjustable parameters: The PBE0 model". In: *The Journal of Chemical Physics* 110.13 (Apr. 1999), pp. 6158–6170. ISSN: 0021-9606. DOI: 10.1063/1.478522.

[163] Narbe Mardirossian and Martin Head-Gordon. " B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy". In: *Phys. Chem. Chem. Phys.* 16 (21 2014), pp. 9904–9924. DOI: 10.1039/C3CP54374A.

[164] W. J. Hehre, R. Ditchfield, and J. A. Pople. "Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules". In: *J. Chem. Phys.* 56 (1972), pp. 2257–2261. DOI: 10.1063/1.1677527.

[165] R. Ditchfield, W. J. Hehre, and J. A. Pople. "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules". In: *J. Chem. Phys.* 54 (1971), pp. 724–728. DOI: 10.1063/1.1674902.

[166] Timothy Clark et al. "Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F". In: *J. Comput. Chem.* 4 (1983), pp. 294–301. DOI: 10.1002/jcc.540040303.

# Chapter 5

# Appendix 1: A Self Consistent Field Formulation of Excited State Mean Field Theory

## 5.1 Mathematical Detail

First, let's rewrite the ESMF ansatz in Eq. (2.10) as

$$|\Psi_{ESMF}\rangle = \sum_{ia} t_{ia} a_{\uparrow}^{\dagger} i_{\uparrow} |\Phi_A\rangle + t_{ia} a_{\downarrow}^{\dagger} i_{\downarrow} |\Phi_A\rangle \tag{5.1}$$

where $\Phi_A$ refers to the Aufbau determinant and $a_{\uparrow}^{\dagger}$ and $i_{\uparrow}^{\dagger}$ are $\uparrow$-spin creation operators for virtual and occupied orbitals, respectively, in the ESMF molecular orbital basis. We define the Hamiltonian in chemist's notation as

$$\begin{aligned} \hat{H} = & \sum_{pq} h_{pq}^{(MO)} \big( p_{\uparrow}^{\dagger} q_{\uparrow} + p_{\downarrow}^{\dagger} q_{\downarrow} \big) \\ & + \frac{1}{2} \sum_{pqrs} \{pq|rs\} \big( p_{\uparrow}^{\dagger} r_{\uparrow}^{\dagger} s_{\uparrow} q_{\uparrow} + p_{\downarrow}^{\dagger} r_{\downarrow}^{\dagger} s_{\downarrow} q_{\downarrow} \\ & \qquad + p_{\uparrow}^{\dagger} r_{\downarrow}^{\dagger} s_{\downarrow} q_{\uparrow} + p_{\downarrow}^{\dagger} r_{\uparrow}^{\dagger} s_{\uparrow} q_{\downarrow} \big) \end{aligned} \tag{5.2}$$

where $\boldsymbol{h}^{(MO)}$ and $\{pq|rs\}$ are the one- and two-electron integrals in the ESMF orbital basis. The ESMF energy is

$$E_{ESMF} = \langle \Psi_{ESMF} | \hat{H} | \Psi_{ESMF} \rangle. \tag{5.3}$$

From here on out, we adopt a summation convention in which a sum is implied over any index that appears more than once in the same term. Using this convention, and working

through the second-quantized algebra, the ESMF energy becomes

$$\begin{aligned} E_{ESMF} = \; & 2\, t_{ia} t_{ib} h^{(MO)}_{ab} - 2\, t_{ia} t_{ja} h^{(MO)}_{ij} + 4\, t_{ia} t_{ia} h^{(MO)}_{kk} \\ & + 4\, t_{ia} t_{ib} \{ab|kk\} - 2\, t_{ia} t_{ib} \{ak|bk\} \\ & - 4\, t_{ia} t_{ja} \{ij|kk\} + 2\, t_{ia} t_{ja} \{ik|jk\} \\ & + 4\, t_{ia} t_{ia} \{kk|ll\} - 2\, t_{ia} t_{ia} \{kl|kl\} \\ & + 4\, t_{ia} t_{jb} \{ia|jb\} - 2\, t_{ia} t_{jb} \{ij|ab\} \end{aligned} \tag{5.4}$$

where $a, b, c, d$ are virtual orbitals, $i, j, k, l, m$ are occupied orbitals, and we will use $p, q, r, s, w, x, y, z$ for general orbitals. Note the pattern in the terms: first, the one electron integrals are contracted and summed over, then the two electron integrals. The two electron integral terms come in pairs, one part of each pair corresponding to the conventional coulomb operator and the other to the conventional exchange operator. As in RHF theory, the coulomb terms have an extra factor of two compared to the exchange terms, which is what produces the general pattern of $2J - K$ in the final mean field operators.

To derive Eq. (2.15), we start with the Lagrangian in Eq. (2.14), take derivatives with respect to the elements of the orbital coefficient matrix $\boldsymbol{C}$, and then set these derivatives equal to zero. Derivatives of the Lagrange multiplier term in Eq. (2.14) lead to the right hand side of Eq. (2.15) in the same way that they do for the HF Roothaan equation, and so we will not work through them explicitly. For the derivatives of the ESMF energy, we need to remember that the one- and two-electron integrals in the ESMF molecular orbital basis are transformed from the atomic orbital basis as

$$h^{(MO)}_{pq} = C_{rp} h_{rs} C_{sq} \tag{5.5}$$

$$\{pq|rs\} = C_{wp} C_{xq} C_{yr} C_{zs} (wx|yz) \tag{5.6}$$

where $\boldsymbol{h}$ and $(wx|yz)$ are the one- and two-electron integrals (again in chemists' notation) in the atomic orbital basis. Noting that $\boldsymbol{C}$ contains an "occupied" block $C^{(o)}$ (the first $n_o$ columns) and a "virtual" block $C^{(v)}$ (the remaining columns), we can consider the derivatives with respect to the elements of these blocks separately. As an example, the occupied block derivative of the first two-electron integral term in Eq. (5.4) is (here $m$ is an occupied index and $x$ is a general index)

$$\frac{\partial}{\partial C_{xm}} \Big( t_{ia} t_{ib} \{ab|kk\} \Big) = \; 2\, t_{ia} t_{ib} C_{pa} C_{qb} C_{sm} (pq|xs). \tag{5.7}$$

The analogous virtual block derivative is

$$\frac{\partial}{\partial C_{xc}} \Big( t_{ia} t_{ib} \{ab|kk\} \Big) = 2\, t_{ic} t_{ia} C_{qa} C_{rk} C_{sk} (xq|rs). \tag{5.8}$$

We now define

$$\mathcal{A}^{coul}_{pq} = C_{rk}C_{sk}(pq|rs) \tag{5.9}$$

$$\mathcal{B}^{coul}_{pq} = C_{sb}t_{ib}t_{ia}C_{ra}(pq|rs) \tag{5.10}$$

$$\mathcal{C}^{coul}_{pq} = C_{ri}t_{ia}t_{ja}C_{sj}(pq|rs) \tag{5.11}$$

$$\mathcal{D}^{coul}_{pq} = C_{ri}t_{ia}C_{sa}(pq|rs) \tag{5.12}$$

as well as $\mathcal{A}^{exch}$, $\mathcal{B}^{exch}$, $\mathcal{C}^{exch}$, and $\mathcal{D}^{exch}$, which are the same except for having $q$ and $r$ swapped in the two-electron integral. With these definitions, we can combine Eqs. (5.7) and (5.8) and write the derivative in matrix form.

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{C}}\Big(t_{ia}t_{ib}\{ab|kk\}\Big) =& \left[\begin{array}{c|c} 2\,\mathcal{B}^{coul}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right] \\ &+ \left[\begin{array}{c|c} \mathbf{0} & 2\,\mathcal{A}^{coul}\boldsymbol{C}^{(v)}\boldsymbol{t}^T\boldsymbol{t} \end{array}\right]\end{aligned} \tag{5.13}$$

where on the right hand side we have placed a vertical bar to separate the occupied and virtual blocks of the matrices. Note that we have kept the two matrices on the right hand side here separate, as they will end up contributing to different terms within Eq. (2.15), which we enumerate as follows.

$$\text{Term 1} \quad \big(\boldsymbol{h} + \boldsymbol{W}[\boldsymbol{A}]\big)C\gamma^{(MO)} \tag{5.14}$$

$$\text{Term 2} \quad \boldsymbol{W}[\boldsymbol{D}]CA^{(MO)} \tag{5.15}$$

$$\text{Term 3} \quad \boldsymbol{W}[\boldsymbol{T}]C(T^{(MO)})^T \tag{5.16}$$

$$\text{Term 4} \quad (\boldsymbol{W}[\boldsymbol{T}])^T CT^{(MO)} \tag{5.17}$$

We will explicitly work through to Terms 1 and 2. Terms 3 and 4 are derived similarly. Now, each of the first nine terms in Eq. (5.4) makes a contribution to Terms 1 or 2. We have already worked out the $\boldsymbol{C}$-derivatives for one of these terms (the fourth one) in Eq. (5.13).

All the others are listed here. Note that the last two derivatives do not make a contribution to Terms 1 or 2, but are listed for completeness.

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ib}h^{(MO)}_{ab}\right) = \left[\begin{array}{c|c} \mathbf{0} & 2\,\boldsymbol{h}\boldsymbol{C}^{(v)}\boldsymbol{t}^T\boldsymbol{t} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ja}h^{(MO)}_{ij}\right) = \left[\begin{array}{c|c} 2\,\boldsymbol{h}\boldsymbol{C}^{(o)}\boldsymbol{t}\boldsymbol{t}^T & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ia}h^{(MO)}_{kk}\right) = \left[\begin{array}{c|c} \boldsymbol{h}\ \boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ib}\{ak|bk\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{B}^{exch}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c} \mathbf{0} & 2\,\mathcal{A}^{exch}\boldsymbol{C}^{(v)}\boldsymbol{t}^T\boldsymbol{t} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ja}\{ij|kk\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{C}^{coul}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c} 2\,\mathcal{A}^{coul}\boldsymbol{C}^{(o)}\boldsymbol{t}\boldsymbol{t}^T & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ja}\{ik|jk\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{C}^{exch}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c} 2\,\mathcal{A}^{exch}\boldsymbol{C}^{(o)}\boldsymbol{t}\boldsymbol{t}^T & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ia}\{kk|ll\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{A}^{coul}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{ia}\{kl|kl\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{A}^{exch}\boldsymbol{C}^{(o)} & \mathbf{0} \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{jb}\{ia|jb\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{D}^{coul}\boldsymbol{C}^{(o)}\boldsymbol{t} & 2\,\mathcal{D}^{coul}\boldsymbol{C}^{(v)}\,\boldsymbol{t}^T \end{array}\right]$$

$$\frac{\partial}{\partial \boldsymbol{C}}\left(t_{ia}t_{jb}\{ij|ab\}\right) = \left[\begin{array}{c|c} 2\,\mathcal{D}^{exch^T}\boldsymbol{C}^{(o)}\boldsymbol{t} & 2\,\mathcal{D}^{exch}\boldsymbol{C}^{(v)}\,\boldsymbol{t}^T \end{array}\right]$$

Using these $\boldsymbol{C}$-derivatives, we get to Term 1 by adding up all the pieces from the derivatives of the terms in Eq. (5.4) that involve $\boldsymbol{h}$, $\mathcal{A}^{coul}$, or $\mathcal{A}^{exch}$. If we make the definition $\mathcal{A} = 2\mathcal{A}^{coul} - \mathcal{A}^{exch}$, then this addition comes out to

$$4\left[\begin{array}{c|c} (\boldsymbol{h}+\mathcal{A})\boldsymbol{C}^{(o)}(\boldsymbol{I}^{(o)} - \boldsymbol{t}\boldsymbol{t}^T) & (\boldsymbol{h}+\mathcal{A})\boldsymbol{C}^{(v)}\boldsymbol{t}^T\boldsymbol{t} \end{array}\right] \tag{5.18}$$

where $I^{(o)}$ is the first $n_o$ columns of $\boldsymbol{I}_o$ (which was defined immediately after Eq. (2.3)). This can be rearranged as

$$4\,(\boldsymbol{h}+\mathcal{A})\;\left[\begin{array}{c|c}\boldsymbol{C}^{(o)} & \boldsymbol{C}^{(v)}\end{array}\right]\;\left[\begin{array}{c|c}\boldsymbol{I}-\boldsymbol{t}\,\boldsymbol{t}^T & \boldsymbol{0}\\ \hline \boldsymbol{0} & \boldsymbol{t}^T\boldsymbol{t}\end{array}\right] \tag{5.19}$$

Now, use the fact that

$$\begin{aligned}\mathcal{A}_{pq} &= 2\mathcal{A}^{coul}_{pq} - \mathcal{A}^{exch}_{pq}\\ &= C_{rk}C_{sk}\big[2(pq|rs) - (pr|qs)\big]\\ &= A_{rs}\big[2(pq|rs) - (pr|qs)\big]\\ &= W[A]_{pq}\end{aligned}$$

to rewrite Eq. (5.19) as

$$4\left(\boldsymbol{h}+\boldsymbol{W}[\boldsymbol{A}]\right)\boldsymbol{C}\,\gamma^{(MO)}$$

which is just Term 1 multiplied by 4. This factor of 4 cancels with the factor of 4 that appears in front of the $\boldsymbol{SC\epsilon}$ term when differentiating the Lagrange multiplier term from Eq. (2.14), thus delivering Term 1 of Eq. (2.15). Term 2 is gotten similarly by collecting the terms that involve $\mathcal{B}^{coul}$, $\mathcal{B}^{exch}$, $\mathcal{C}^{coul}$, or $\mathcal{C}^{exch}$. Defining $\mathcal{B}$ and $\mathcal{C}$ analogous to $\mathcal{A}$, we get:

$$4(\mathcal{B}-\mathcal{C})\left[\begin{array}{c|c}\boldsymbol{C}^{(o)} & \boldsymbol{0}\end{array}\right] \tag{5.20}$$

However, recall that

$$\begin{aligned}(\mathcal{B}-\mathcal{C})_{pq} &= 2\mathcal{B}^{coul}_{pq} - \mathcal{B}^{exch}_{pq} - 2\mathcal{C}^{coul}_{pq} + \mathcal{C}^{exch}_{pq}\\ &= C_{sb}t_{ib}t_{ia}C_{ra}\big[2(pq|rs) - (pr|qs)\big]\\ &\quad - C_{ri}t_{ia}t_{ja}C_{sj}\big[2(pq|rs) - (pr|qs)\big]\\ &= \left(\left[\begin{array}{c|c}\boldsymbol{C}^{(o)} & \boldsymbol{C}^{(v)}\end{array}\right]\left[\begin{array}{c|c}-\boldsymbol{t}\,\boldsymbol{t}^T & \boldsymbol{0}\\ \hline \boldsymbol{0} & \boldsymbol{t}^T\boldsymbol{t}\end{array}\right]\left[\begin{array}{c|c}\boldsymbol{C}^{(o)} & \boldsymbol{C}^{(v)}\end{array}\right]^T\right)_{rs}\\ &\qquad \cdot\big[2(pq|rs) - (pr|qs)\big]\\ &= W[D]_{pq}\end{aligned}$$

Then, Eq. (5.20) can be written as

$$4(\mathcal{B}-\mathcal{C})\left[\begin{array}{c|c}\boldsymbol{C}^{(o)} & \boldsymbol{0}\end{array}\right] = 4W[D]CA^{(MO)}$$

which, again cancelling the factor of 4, is Term 2. Terms 3 and 4 work similarly, this time collecting the terms involving $\mathcal{D}^{coul}$ and $\mathcal{D}^{exch}$.

Note that Eq. (2.11) can be obtained by rearranging Eq. (5.4). This follows from regrouping terms the same way as during the derivation of Eq. (2.15).

The commutator relation given in Eq. (2.16) can be derived from Eq. (2.15), analogous to how Eq. (2.8) is derived from Eq. (2.2). First, multiply the left hand side of Eq. (2.15) by $C^T$ on the left. Then, transpose the left hand side of Eq. (2.15) and multiply $C$ on the right. Finally, take the difference.

$$C^T \left(\text{Eq.(2.15) LHS}\right) - \left(\text{Eq.(2.15) LHS}\right)^T C \tag{5.21}$$

Because $\epsilon$ is symmetric and $\boldsymbol{C}^T\boldsymbol{S}\boldsymbol{C} = I$, the equivalent manipulations of the right hand sides of Eq. (2.15) subtract to give zero. Thus, Eq. (5.21) must be equal to zero, giving us

$$\begin{aligned} C^T F_A C\gamma^{(MO)} - &\ (\gamma^{(MO)})^T C^T F_A^T C \\ + C^T W[D] C A^{(MO)} - &\ (A^{(MO)})^T C^T (W[D])^T C \\ + C^T W[T] C (T^{(MO)})^T - &\ T^{(MO)} C^T (W[T])^T C \\ + C^T (W[T])^T C T^{(MO)} - &\ (T^{(MO)})^T C^T W[T] C = 0. \end{aligned} \tag{5.22}$$

Recall that $F_A$ is the regular RHF Fock matrix, $W[D]$ is a mean-field operator defined on a difference of density matrices, and $\gamma^{(MO)}$, $A^{(MO)}$ are 1-body RDMs, making these matrices symmetric. For terms with $W[T]$ and $T$, the transposes are part of the expression already. Therefore, this simplifies to the commutator relation as given in Eq. (2.16).

Now, starting with Eq. (2.16), which may not be satisfied for the $\boldsymbol{C}$ matrix we have, we seek to make an improvement by rotating $\boldsymbol{C}$ as $\boldsymbol{C}e^{\boldsymbol{X}}$, where $\boldsymbol{X}$ is an anti-symmetric matrix. Assuming that the rotation will be small, we approximate $e^{\boldsymbol{X}} = \boldsymbol{I} + \boldsymbol{X}$. Inserting this approximation into Eq. (2.16) gives

$$\begin{aligned} 0 = &\left[\ (C + CX)^T F_A (C + CX),\ \gamma^{(MO)}\ \right] \\ &+ \left[\ (C + CX)^T W[D] (C + CX),\ A^{(MO)}\ \right] \\ &+ \left[\ (C + CX)^T W[T] (C + CX),\ (T^{(MO)})^T\ \right] \\ &+ \left[\ (C + CX)^T (W[T])^T (C + CX),\ T^{(MO)}\ \right]. \end{aligned} \tag{5.23}$$

Dropping any terms that have more than one power of $\boldsymbol{X}$ in them and rearranging into the form of a linear equation for $\boldsymbol{X}$, we arrive at the working linear equation that we solve with GMRES to find a new $\boldsymbol{X}$ and then update $\boldsymbol{C}$.

$$
\begin{aligned}
& - \left[F_A^{(MO)}, \gamma^{(MO)}\right] - \left[W[D]^{(MO)}, A^{(MO)}\right] \\
& - \left[W[T]^{(MO)}, (T^{(MO)})^T\right] - \left[(W[T]^{(MO)})^T, T^{(MO)}\right] \\
= \quad & \left(F_A^{(MO)} X - X F_A^{(MO)}\right) \gamma^{(MO)} \\
& - \gamma^{(MO)} \left(F_A^{(MO)} X - X F_A^{(MO)}\right) \\
& + \left(W[D]^{(MO)} X - X W[D]^{(MO)}\right) A^{(MO)} \\
& - A^{(MO)} \left(W[D]^{(MO)} X - X W[D]^{(MO)}\right) \\
& + \left(W[T]^{(MO)} X - X W[T]^{(MO)}\right) (T^{(MO)})^T \\
& - (T^{(MO)})^T \left(W[T]^{(MO)} X - X W[T]^{(MO)}\right) \\
& + \left((W[T]^{(MO)})^T X - X (W[T]^{(MO)})^T\right) T^{(MO)} \\
& - T^{(MO)} \left((W[T]^{(MO)})^T X - X (W[T]^{(MO)})^T\right) . \qquad (5.24)
\end{aligned}
$$

Note that, once we use $\boldsymbol{X}$ to update $\boldsymbol{C}$, we reevaluate the mean-field operators $F_A$, $W[D]$, and $W[T]$. Once multiple SCF orbital optimization iterations are complete and the error in Eq. (2.16) is small (as measured for example by by $\Delta$, see next section), we perform a new CIS calculation in the new orbital basis to update $\boldsymbol{t}$, and then start in again on orbital

**Table 5.1:** ESMF details in PYCM charge transfer state.

| Iteration Type | Energy (a.u.) | $\Delta$ (a.u.) | DIIS? |
|---|---|---|---|
| ORB OPT | -571.178433339545 | 3.5554e-01 | NO |
| ORB OPT | -571.246925619201 | 2.0214e-01 | NO |
| ORB OPT | -571.262163759189 | 1.6952e-01 | YES |
| ORB OPT | -571.277606357895 | 2.7521e-02 | YES |
| ORB OPT | -571.278746939584 | 9.0729e-03 | YES |
| ORB OPT | -571.278998781297 | 3.9852e-03 | YES |
| ORB OPT | -571.279081935162 | 2.5653e-03 | YES |
| ORB OPT | -571.279097212426 | 8.7319e-04 | YES |
| ORB OPT | -571.279100065342 | 4.2506e-04 | YES |
| ORB OPT | -571.279100566347 | 1.9735e-04 | YES |
| ORB OPT | -571.279100661748 | 8.8807e-05 | YES |
| CIS | -571.279215755055 | N/A | N/A |
| ORB OPT | -571.279215755013 | 4.9659e-04 | NO |
| ORB OPT | -571.279216054833 | 2.4730e-04 | NO |
| ORB OPT | -571.279216098757 | 1.8869e-04 | YES |
| ORB OPT | -571.279216133376 | 5.1892e-05 | YES |
| CIS | -571.279216139390 | N/A | N/A |

optimization via our linear equation. As a final note, we use a diagonal preconditioner to approximate the inverse of the linear transformation and accelerate the GMRES solution of the linear equation. For elements of $\boldsymbol{X}$ in the occupied-occupied or virtual-virtual blocks, the preconditioner's diagonal element is just one. For elements of $\boldsymbol{X}$ in the occupied-virtual or virtual-occupied blocks, i.e. $X_{ia}$ or $X_{ai}$, the preconditioner's diagonal element is

$$\left( \left[ F_A^{(MO)} \right]_{aa} - \left[ F_A^{(MO)} \right]_{ii} \right)^{-1} \tag{5.25}$$

## 5.2 Charge Transfer Iteration Details

For the charge transfer optimization in PYCM, we show in Table 5.1 the total energy as well as $\Delta$, the Frobenius norm of the error in the commutator matrix equation in the molecular orbital basis, at each stage of ESMF's SCF optimization. Each "ORB OPT" iteration corresponds to one solution of the linear equation for X. The "DIIS?" column labels which orbital optimization steps employed DIIS. After optimal orbitals for the current $\boldsymbol{t}$ coefficients are found, a CIS calculation is performed in the new molecular orbital basis in order to update $\boldsymbol{t}$. Note that the initial guess in this case was to set the $\boldsymbol{C}$ and $\boldsymbol{t}$ matrices to correspond to a HOMO→LUMO transition in the RHF orbital basis, which corresponds to the charge transfer state in question. The orbitals were then optimized by our SCF procedure, after which a CIS calculation updated $\boldsymbol{t}$, after which the orbitals were optimized again, after which another CIS calculation shows that the energy has converged.

## 5.3 Geometries

Molecular geometries (some in Angstrom and some in Bohr, see table labels) are given in the following tables. Note that the PYCM geometry is on the next page.

| Molecule: Water | | Units: Angstrom | |
|---|---|---|---|
| O | 0.0000000 | 0.0000000 | 0.1157190 |
| H | 0.0000000 | 0.7487850 | -0.4628770 |
| H | 0.0000000 | -0.7487850 | -0.4628770 |

| Molecule: Ethylene | | Units: Angstrom | |
|---|---|---|---|
| C | 0.0000000000 | 0.0000000000 | 0.6727698502 |
| C | 0.0000000000 | 0.0000000000 | -0.6727698502 |
| H | 0.0000000000 | -0.9347680531 | 1.2426974978 |
| H | 0.0000000000 | 0.9347680530 | 1.2426974980 |
| H | 0.0000000000 | -0.9347680532 | -1.2426974978 |
| H | 0.0000000000 | 0.9347680530 | -1.2426974980 |

| Molecule: Formaldehyde | | Units: Angstrom | |
|---|---|---|---|
| C | -0.0910041349 | 0.1032665042 | -0.6906844847 |
| O | -0.7161280585 | 0.7732151716 | 0.1085248268 |
| H | -0.2415602237 | 0.1918899655 | -1.7934092839 |
| H | 0.6751090789 | -0.6449677992 | -0.3749139325 |

| Molecule: Toluene | | Units: Angstrom | |
|---|---|---|---|
| C | 0.0865862990 | 0.7410453762 | 0.0688914923 |
| C | -0.1345668535 | -0.6531059507 | 0.0602786905 |
| C | -1.4505649626 | -1.1603576134 | -0.0093490679 |
| C | -2.5405851883 | -0.2644802090 | -0.0661246831 |
| C | -2.3124601785 | 1.1285606766 | -0.0571533440 |
| C | -0.9977830949 | 1.6502417205 | 0.0134992344 |
| C | -0.7535284502 | 3.1540596435 | -0.0111519475 |
| H | 1.1079021258 | 1.1261067270 | 0.1292876906 |
| H | 0.7134665908 | -1.3382132205 | 0.1093728295 |
| H | -1.6243218250 | -2.2374292057 | -0.0134283434 |
| H | -3.5611922298 | -0.6477509085 | -0.1151906890 |
| H | -3.1617587236 | 1.8157708718 | -0.0950213452 |
| H | -0.6369699217 | 3.5271815420 | -1.0399480108 |
| H | 0.1595927745 | 3.4107593481 | 0.5443343618 |
| H | -1.5935598619 | 3.6938717021 | 0.4484031317 |

| Molecule: PYCM | | Units: Bohr | |
|---|---|---|---|
| C | 7.48914118998055 | -2.05146035486422 | -1.28980829449057 |
| C | 5.98770398368311 | 0.04915735478248 | -0.06865458045111 |
| C | 7.61081607331134 | 2.27669871460533 | 0.67560530219819 |
| C | 3.46706532904223 | -0.09116628967473 | 0.30637708021880 |
| C | 1.92647158203893 | -2.41958847844936 | -0.43805415580348 |
| C | -0.75455162629890 | -2.43014371114785 | 0.69884575028668 |
| C | -2.18486749739264 | -0.06198337612267 | 0.06973201498199 |
| C | -0.64398578220618 | 2.33392846576440 | 0.05840028021288 |
| C | 1.88958084771184 | 2.01941654524793 | 1.44673369015458 |
| C | -4.70904737992718 | -0.06309418170513 | -0.46180929064412 |
| C | -6.17576878134698 | -2.34798184267961 | -0.48999098912292 |
| N | -7.35364877360203 | -4.20620382240250 | -0.50945688691185 |
| C | -6.03132613865057 | 2.24217192107952 | -1.01894902061744 |
| N | -7.06517180189986 | 4.13250579177077 | -1.46439337259584 |
| H | 8.54937612247998 | -1.31872561024836 | -2.93252209992581 |
| H | 6.33486664187961 | -3.64092303587743 | -1.95285217876895 |
| H | 8.91885629767513 | -2.80382984977844 | 0.03384049526409 |
| H | 9.17826984585370 | 1.64141995115378 | 1.89968164549696 |
| H | 6.59508025455693 | 3.76874545170370 | 1.69452668746078 |
| H | 8.50253869488760 | 3.13726625417923 | -1.00559579441233 |
| H | 2.88064720049037 | -4.16472134045315 | 0.16752333485011 |
| H | 1.76036799956004 | -2.54566120290402 | -2.51696904366760 |
| H | -0.60805363262071 | -2.53086096937472 | 2.78367809044349 |
| H | -1.80585658070689 | -4.12167616649262 | 0.11092207863213 |
| H | -1.74995920022716 | 3.90829005165278 | 0.85174533709869 |
| H | -0.27696758321317 | 2.85096221165140 | -1.93484378989024 |
| H | 1.49786269333475 | 1.62230247974489 | 3.46233841913685 |
| H | 2.89384261982421 | 3.83106450894856 | 1.40264023881446 |

# Chapter 6

# Appendix 2: Exploring Mulliken Population Changes upon Excitation in Explicitly Solvated Systems

## 6.1 Geometries

All solvated geometries were generated by CREST and are given in Angstrom.

| Molecule: 4-cyano pyridine | | | |
|---|---|---|---|
| C | -1.49744 | -1.138724 | 0.000002 |
| C | -0.11206 | -1.19872 | -0.000001 |
| N | -2.18409 | 0.00000 | 0.000004 |
| C | 0.59178 | -0.00000 | -0.000001 |
| C | 2.02889 | -0.00000 | -0.000004 |
| C | -0.11206 | 1.19872 | 0.000001 |
| N | 3.17532 | 0.00000 | -0.000006 |
| C | -1.49743 | 1.13872 | 0.000004 |
| H | -2.08043 | -2.05147 | 0.000002 |
| H | 0.40608 | -2.14628 | -0.000002 |
| H | 0.40609 | 2.14628 | 0.000001 |
| H | -2.08042 | 2.05147 | 0.000005 |

| Molecule: 4-cyano pyridine with 4 water molecules | | | |
|---|---|---|---|
| C | -1.801128098 | -1.290734633 | -0.4654940286 |
| C | -0.4851121286 | -1.058933353 | -0.1063815834 |
| N | -2.661347471 | -0.3274273664 | -0.7534783871 |
| C | -0.05326584067 | 0.2656315468 | -0.04781416424 |
| C | 1.275478818 | 0.5690617655 | 0.3219550328 |
| C | -0.955849155 | 1.287590078 | -0.3378175283 |
| N | 2.350379598 | 0.8092194752 | 0.6408307312 |
| C | -2.246415263 | 0.926795677 | -0.68377957 |
| H | -2.189798909 | -2.298584413 | -0.531815833 |
| H | 0.1860781161 | -1.872252536 | 0.1196867883 |
| H | -0.6531944963 | 2.321799272 | -0.2950239895 |
| H | -2.988213791 | 1.677789432 | -0.9227792934 |
| O | -0.08264179493 | 0.2079520454 | 2.978208179 |
| H | 0.4938381521 | -0.5796100264 | 2.969122382 |
| H | -0.6554565861 | 0.1056958461 | 3.740313751 |
| O | 2.000049501 | -1.596341709 | 3.089508798 |
| H | 2.517324175 | -1.869400943 | 2.330003306 |
| H | 2.430329161 | -0.7824043776 | 3.424170679 |
| O | 2.632862429 | 1.020442947 | 3.498119612 |
| H | 2.981817551 | 1.225741854 | 2.622616075 |
| H | 1.675895317 | 1.144574815 | 3.427538465 |
| O | -0.9799442892 | -0.3269699952 | -3.42529592 |
| H | -0.9539056324 | -0.4215284982 | -4.380017342 |
| H | -1.892680225 | -0.4580071409 | -3.154803188 |

| Molecule: formaldehyde with 8 waters | | | |
|---|---|---|---|
| H | -0.26208813043283 | 1.52324614836254 | -1.44055149046130 |
| H | -0.96531439110153 | -0.17757703807398 | -1.35476471452980 |
| C | -0.08425203974286 | 0.46237891114631 | -1.22190908685695 |
| O | 1.04969753994279 | -0.01716777045278 | -1.29549206383906 |
| O | 2.80743859152177 | 0.47266954043778 | 2.23013473386229 |
| H | 2.29202936847883 | -0.32803172728579 | 2.06048468382088 |
| H | 3.21224817462003 | 0.70616200514085 | 1.38051427430079 |
| O | -3.00918953173046 | -0.09451696122368 | 1.25756841672731 |
| H | -2.18021862530018 | 0.34291700867676 | 0.99366461265194 |
| H | -2.96790615146923 | -0.19641081009137 | 2.21033520652050 |
| O | 1.06466672939600 | -1.42741915569149 | 1.07754430681112 |
| H | 0.51363364353834 | -2.24005402482116 | 1.10035773632174 |
| H | 1.32746960215150 | -1.30583245097695 | 0.15073305926862 |
| O | 1.20138893004835 | 2.48537968952001 | 1.49927657041952 |
| H | 1.74160409113634 | 1.87369755920882 | 2.04870430280640 |
| H | 1.77465258697489 | 2.67453705823828 | 0.74562638359037 |
| O | -0.44315345966638 | 0.68075014419933 | 0.65463597487236 |
| H | 0.11617274838525 | 1.46594223030995 | 0.97403351347724 |
| H | 0.04810674784033 | -0.12904042082709 | 0.98031188160275 |
| O | 3.07958166329960 | 1.48360553972951 | -0.39201237702588 |
| H | 3.78808450251042 | 1.49942555211271 | -1.03535846421318 |
| H | 2.36227149830573 | 0.93419501536525 | -0.78646865940856 |
| O | -2.81487626938565 | -2.37631276180632 | -0.17798353682149 |
| H | -3.65426157351825 | -2.83475115340114 | -0.11346847000282 |
| H | -2.93172550568819 | -1.52454941727453 | 0.29519693300871 |
| O | -0.58782394451866 | -3.52264108584505 | 0.82271058742592 |
| H | -1.40641426824753 | -3.11840562236029 | 0.46157750596076 |
| H | -0.31663167482658 | -4.18939899737191 | 0.19016597833379 |

| Molecule: formaldehyde with 12 waters | | | |
|---|---|---|---|
| H | -1.16269866942290 | 0.77631511092895 | -0.36316067486297 |
| H | -1.13238792574785 | -0.99070149660763 | 0.21850662170440 |
| C | -0.57975553117158 | -0.07801934456137 | -0.01767948496467 |
| O | 0.60982780775428 | 0.00361654737168 | 0.18486638707656 |
| O | 1.83110782618584 | -1.37829632519190 | -2.16007706908093 |
| H | 1.94394562338012 | -1.89840084549036 | -1.34652555373311 |
| H | 2.16659956653851 | -0.49028212056156 | -1.94005474333315 |
| O | -3.04813479779155 | -1.73822370408430 | -1.05150199966519 |
| H | -2.57499543784558 | -2.47230338584032 | -0.63388189010981 |
| H | -2.40122018098032 | -1.37138883303994 | -1.67355532147404 |
| O | 1.81685292265675 | -2.36402237914590 | 0.45501663537034 |
| H | 2.61615007475824 | -2.34949327791775 | 0.98354818053873 |
| H | 1.44529914120449 | -1.45585054127107 | 0.53014828299043 |
| O | -0.71253862359885 | -0.72230925077242 | -2.24476414359052 |
| H | 0.22642243783891 | -1.05526624938753 | -2.27429917039334 |
| H | -0.70693926936362 | 0.10388197208886 | -2.73175933304922 |
| O | 3.13978947823343 | 0.40982813764292 | 1.26125076072088 |
| H | 2.89764390540416 | 0.57570490717349 | 0.33082610856288 |
| H | 3.88026006104896 | 0.98645958782117 | 1.45437848196078 |
| O | -3.66945793447644 | 0.36793090125472 | 0.50374149408856 |
| H | -4.62011864761215 | 0.36452493197348 | 0.62411174476769 |
| H | -3.46774674916890 | -0.40613059338687 | -0.07667107890937 |
| O | 2.59464669906497 | 1.18029558045133 | -1.38082359465105 |
| H | 3.17889967838460 | 1.75274757465412 | -1.87935483695947 |
| H | 1.83111067933097 | 1.74901090421173 | -1.10410502885324 |
| O | -1.57482421072450 | 0.74882823435084 | 2.30905456602972 |
| H | -2.36544351824897 | 0.62630756312049 | 1.75573953432519 |
| H | -1.56131922315078 | 0.02361838791197 | 2.93535796748758 |
| O | -0.75093596914161 | -3.33156606269529 | -0.37828571677333 |
| H | 0.09243494665995 | -3.16229034948223 | 0.06589601189635 |
| H | -0.53710844091085 | -3.36696128824453 | -1.31378122224078 |
| O | 1.00776188871252 | 2.11428977962467 | 2.04947647661754 |
| H | 1.66561552755088 | 1.41439976978572 | 1.89117178029758 |
| H | 0.16745140028172 | 1.66216864655864 | 2.19641281789235 |
| O | 0.76249615385269 | 2.91238144144145 | -0.53296617303457 |
| H | 0.89097437703151 | 2.74640224952556 | 0.42731584275961 |
| H | -0.19449321134686 | 2.94037863829591 | -0.65174303077249 |
| O | -2.13562312378373 | 2.78213472946622 | 0.00197480718130 |
| H | -1.86734928080620 | 2.95347319054349 | 0.90888957408378 |
| H | -2.80972927188136 | 2.09026549518273 | 0.05398696739749 |

| Molecule: PYCM with 4 waters | | | |
|---|---|---|---|
| C | 4.01488288 | -1.558054041 | 0.3665908313 |
| C | 3.150015763 | -0.4486937012 | -0.1730868674 |
| C | 3.762898725 | 0.3061556324 | -1.320537643 |
| C | 1.955165273 | -0.1527549779 | 0.3396047192 |
| C | 1.35380166 | -0.8890724934 | 1.507412913 |
| C | -0.05167370397 | -1.377981695 | 1.156762174 |
| C | -0.8874069876 | -0.2954934183 | 0.5844950186 |
| C | -0.2591061265 | 1.050394911 | 0.5341641473 |
| C | 1.078485874 | 0.9487609694 | -0.2105883583 |
| C | -2.128811587 | -0.5305372259 | 0.08615794019 |
| C | -2.725174927 | -1.811975523 | 0.08784854881 |
| N | -3.227478802 | -2.84220879 | 0.06834669772 |
| C | -2.88989431 | 0.4717380645 | -0.5392457119 |
| N | -3.492007598 | 1.277682942 | -1.096103325 |
| H | 4.955008188 | -1.600296903 | -0.1765527559 |
| H | 3.517079479 | -2.521133883 | 0.260053332 |
| H | 4.241092942 | -1.406683393 | 1.420839628 |
| H | 4.159916501 | -0.3876287276 | -2.061468728 |
| H | 4.600348099 | 0.9023443971 | -0.9557518448 |
| H | 3.060804689 | 0.974648103 | -1.809003009 |
| H | 1.314330544 | -0.2169342723 | 2.371205371 |
| H | 1.96073154 | -1.746704135 | 1.78442198 |
| H | -0.5399862318 | -1.803790007 | 2.038962374 |
| H | 0.02985771023 | -2.201454934 | 0.4343482183 |
| H | -0.07832493414 | 1.391285503 | 1.55805939 |
| H | -0.903981754 | 1.779409669 | 0.04085056213 |
| H | 1.590278856 | 1.912130146 | -0.1443254008 |
| H | 0.8554264972 | 0.7637145599 | -1.264478995 |
| O | -0.850354741 | -0.2474718796 | -2.889468459 |
| H | -1.128924352 | 0.656866941 | -3.120067675 |
| H | -1.540050179 | -0.8289140699 | -3.215835871 |
| O | -2.092991677 | 2.169669035 | -3.373059023 |
| H | -2.6053965 | 2.466458809 | -4.12727798 |
| H | -2.710432326 | 2.094624076 | -2.626848932 |
| O | 1.409377006 | -1.853951672 | -2.728768982 |
| H | 2.039807088 | -1.374696377 | -2.184725746 |
| H | 0.6459826137 | -1.262992045 | -2.839517971 |
| O | 0.5314044435 | -3.652106867 | -0.8749576694 |
| H | -0.07391957999 | -4.282254908 | -1.268532212 |

| | | | |
|---|---|---|---|
| H | 0.8839989547 | -3.109919657 | -1.608556003 |

| Molecule: PYCM with 10 waters | | | |
|---|---|---|---|
| C | 4.080816988 | -1.100115929 | -0.2671534111 |
| C | 3.059566859 | -0.004744586007 | -0.1640159479 |
| C | 3.612674582 | 1.371771349 | -0.4256703721 |
| C | 1.769567026 | -0.1703039039 | 0.139580355 |
| C | 1.093991372 | -1.49038409 | 0.3939567508 |
| C | -0.2343044469 | -1.561135348 | -0.3583754869 |
| C | -1.141515702 | -0.4477791169 | 0.008841788208 |
| C | -0.5400308404 | 0.6697530316 | 0.7700130031 |
| C | 0.8453110043 | 1.018208427 | 0.2223527236 |
| C | -2.451837853 | -0.4516309518 | -0.3569908345 |
| C | -3.009258706 | -1.505344523 | -1.10839661 |
| N | -3.446726409 | -2.362754234 | -1.733117529 |
| C | -3.334816 | 0.5847465549 | -0.02572321401 |
| N | -4.082912319 | 1.407518759 | 0.2647448859 |
| H | 4.555573929 | -1.074979018 | -1.248232699 |
| H | 3.664519185 | -2.089654027 | -0.1121484933 |
| H | 4.860001217 | -0.9402277762 | 0.4782347536 |
| H | 4.657602681 | 1.315501894 | -0.7171880042 |
| H | 3.540936648 | 1.995006899 | 0.4659226168 |
| H | 3.059674818 | 1.871643811 | -1.220110293 |
| H | 0.9114584225 | -1.602599678 | 1.467679517 |
| H | 1.7134221 | -2.326178792 | 0.07587595355 |
| H | -0.7195007432 | -2.52737135 | -0.1961545426 |
| H | -0.04052220415 | -1.495108189 | -1.434526019 |
| H | -0.4562682369 | 0.3605511929 | 1.817256888 |
| H | -1.179121764 | 1.551578033 | 0.74329061 |
| H | 1.284717488 | 1.790578452 | 0.8572686987 |
| H | 0.7200398501 | 1.456024474 | -0.7723695618 |
| O | -0.8541258791 | 0.1246054986 | -3.056998763 |
| H | 0.1083312126 | 0.09958479351 | -3.217167381 |
| H | -1.052830488 | 1.019848494 | -2.748483555 |
| O | -1.517214497 | -2.325693825 | -4.047900656 |
| H | -2.258222732 | -2.624572099 | -3.510135532 |
| H | -1.403520123 | -1.380764827 | -3.837951957 |
| O | 1.831574494 | -0.4350317057 | -3.184245239 |
| H | 2.240525921 | -0.352244362 | -2.31769603 |
| H | 1.622183362 | -1.386097872 | -3.288969992 |
| O | -2.052639767 | 0.737135006 | 3.615120191 |
| H | -1.819540177 | 0.70642238 | 4.54484803 |
| H | -3.029083704 | 0.6526120476 | 3.584323738 |

| | | | |
|---|---|---|---|
| O | 0.9754847425 | -3.014734385 | -3.346331621 |
| H | 1.341691843 | -3.521851378 | -4.072090239 |
| H | 0.03883035867 | -2.833394099 | -3.598561881 |
| O | -4.759476567 | 0.7641893987 | 3.49707809 |
| H | -4.913019874 | 1.675730203 | 3.195198162 |
| H | -5.165147063 | 0.1944618806 | 2.840835043 |
| O | -4.907750456 | 3.168062694 | 2.14110612 |
| H | -4.666236165 | 2.663063028 | 1.336710985 |
| H | -5.476441385 | 3.885949123 | 1.860814125 |
| O | -2.180923273 | 3.371649833 | 2.78435463 |
| H | -2.000279902 | 2.447796996 | 3.038128513 |
| H | -3.145189366 | 3.433297329 | 2.706386072 |
| O | -1.342347664 | 4.063577447 | 0.3452029877 |
| H | -0.6308721594 | 4.693350446 | 0.4669973585 |
| H | -1.645813535 | 3.818727883 | 1.250347495 |
| O | -1.4147519 | 2.655524957 | -2.000534397 |
| H | -1.365352629 | 3.115134203 | -1.142181381 |
| H | -2.10149944 | 3.102350662 | -2.497250579 |

| Molecule: PYCM with 20 waters | | | |
|---|---|---|---|
| C | 4.13850615504417 | -1.33284491194929 | -0.15010572959486 |
| C | 3.25101651055397 | -0.13021595275574 | -0.31026328716862 |
| C | 3.91727330143164 | 1.05430530242872 | -0.95800846535663 |
| C | 1.98619821732314 | -0.04234545113392 | 0.10540546874085 |
| C | 1.19383824964747 | -1.14094530940111 | 0.74995581681216 |
| C | -0.16639813038528 | -1.26549636578067 | 0.05916552581602 |
| C | -0.95028367420526 | -0.02004073051146 | 0.18290259797686 |
| C | -0.19043221831510 | 1.19611307022487 | 0.52321749954555 |
| C | 1.20801671401723 | 1.23176603338170 | -0.09129081184978 |
| C | -2.30633153275588 | -0.00493413573503 | 0.04213501354185 |
| C | -3.03765711166430 | -1.19052592743294 | -0.10343885565704 |
| N | -3.62840436247943 | -2.17696971065598 | -0.15129735863001 |
| C | -3.05604179199583 | 1.18649613887011 | 0.11389102246169 |
| N | -3.65961940707909 | 2.16082570544558 | 0.18155217086477 |
| H | 4.67427084978300 | -1.53567072530011 | -1.07807424168597 |
| H | 3.59966367044726 | -2.22949418673379 | 0.14034565476456 |
| H | 4.88665606983544 | -1.13042767754264 | 0.61699116114807 |
| H | 4.92369456730810 | 0.80406055515147 | -1.28191240274444 |
| H | 3.98073235062740 | 1.88685825313015 | -0.25775097987798 |
| H | 3.34925790289473 | 1.40063775645676 | -1.82094090121602 |
| H | 1.04351745195394 | -0.90474835487219 | 1.80567962892544 |
| H | 1.70899732762576 | -2.09577333147833 | 0.70104289674386 |
| H | -0.72406898470153 | -2.11668658749896 | 0.44338067640352 |
| H | -0.00287702688847 | -1.45064240644626 | -1.00638967134507 |
| H | -0.08327360566230 | 1.18859708057712 | 1.61353390714711 |
| H | -0.74651979212313 | 2.09422952033767 | 0.26210015676191 |
| H | 1.75420415470457 | 2.06400706378677 | 0.35497083032353 |
| H | 1.12720666714902 | 1.43918712441672 | -1.15827129226663 |
| O | -2.39540863179645 | -1.53865062915854 | -3.47030630109002 |
| H | -3.10887076061239 | -1.00931063002024 | -3.08533713780795 |
| H | -1.62961734461891 | -0.95230190172090 | -3.54659058271202 |
| O | 0.00858966827482 | -0.04578061532063 | -3.23052233039125 |
| H | 0.58364735977838 | 0.72569055556700 | -3.36974927393834 |
| H | 0.59677110903194 | -0.78968536748894 | -3.02220447808196 |
| O | -2.00175905838663 | 1.81010933837159 | -2.57088462155450 |
| H | -1.54998356906521 | 2.56543589509414 | -2.16362021004255 |
| H | -1.31067418911412 | 1.16331720133276 | -2.77662139331967 |
| O | -1.00562948404437 | -3.70946198715759 | -2.63876644213076 |
| H | -0.14149480592715 | -3.36692291105089 | -2.88522759782850 |
| H | -1.64086187189276 | -3.04181780028234 | -2.96380112472441 |

| | | | |
|---|---|---|---|
| O | -0.41253856715069 | 3.92286240956508 | -1.60587312066326 |
| H | -0.21008897065638 | 4.05395029389021 | -0.64108323783834 |
| H | -0.74106371868348 | 4.77090873516136 | -1.90817712340097 |
| O | -2.32190207398710 | 3.57859362239483 | 2.27218292629903 |
| H | -2.34468332849683 | 2.67747122204283 | 2.63136540352755 |
| H | -3.02034846026705 | 3.59469328578920 | 1.60797458100954 |
| O | -1.36172784837051 | -1.65261362621088 | 3.09040328642225 |
| H | -0.89354565584972 | -2.37148116327409 | 2.64260385836555 |
| H | -2.27916356283461 | -1.98025888811801 | 3.18163083514148 |
| O | -5.19290467341859 | 0.04606647568291 | 2.19967568802656 |
| H | -5.43600092386143 | 0.01361146796558 | 1.24634975194441 |
| H | -5.92782839134545 | 0.47419749909588 | 2.64156655767535 |
| O | -5.94477709630853 | -0.21759661538459 | -0.38002919807345 |
| H | -6.01407877925696 | -1.16061294388368 | -0.54884973721424 |
| H | -5.35704751520394 | 0.12825638274579 | -1.07605691155447 |
| O | -4.26258338704458 | 0.38500229424918 | -2.51495775560289 |
| H | -4.78441777863216 | 0.68747405757485 | -3.26058367929492 |
| H | -3.50335719163598 | 1.00553308940269 | -2.45207980802344 |
| O | -2.64848090540609 | 0.89953001936487 | 3.05153962222798 |
| H | -2.20562305892970 | 0.03908294801808 | 3.00858296857843 |
| H | -3.57077746670061 | 0.72300835839978 | 2.81262837805496 |
| O | -3.97980198198955 | -2.37380268812964 | 2.84972342033456 |
| H | -4.46856973273367 | -1.54012292441881 | 2.74307630018970 |
| H | -4.03760249836420 | -2.81693083128547 | 1.99764834633957 |
| O | 0.08179397594676 | 0.84144108953019 | 3.85013855186153 |
| H | -0.76576615150180 | 1.27072206811602 | 4.01145275383373 |
| H | -0.12975595490232 | -0.09657919912718 | 3.75503198749904 |
| O | 1.59934553660706 | 2.94341568272997 | 2.73843299886463 |
| H | 1.86397207013781 | 3.48906236463566 | 3.48019627493195 |
| H | 1.16744731845832 | 2.16136610904262 | 3.12043638019430 |
| O | -0.04559414852723 | 4.37245978264215 | 0.97339417311893 |
| H | 0.63880038899782 | 3.91388337890814 | 1.48862852762670 |
| H | -0.87397199030510 | 4.20564205342136 | 1.46479909379794 |
| O | 1.64990760104538 | -2.18780971818301 | -2.40278108485771 |
| H | 1.57121036707849 | -2.91643905152953 | -1.75253036944704 |
| H | 2.43140302256478 | -1.69748424433517 | -2.13503051200718 |
| O | 1.23383362304618 | 2.44331466582074 | -3.27941084258912 |
| H | 1.32514978312110 | 2.94333588238373 | -4.09161010758928 |
| H | 0.71063696400890 | 2.99711057090307 | -2.66784589017420 |
| O | -2.19507838224383 | -4.57932507827314 | -0.35546666489021 |
| H | -1.74868290365406 | -4.37792292153607 | -1.20166412829912 |
| H | -2.94807374226602 | -3.97278195425944 | -0.33129715081419 |

| | | | |
|---|---|---|---|
| O | -0.51928266752831 | -4.03691964700294 | 1.60457800373668 |
| H | -0.47149819481513 | -4.77315880883380 | 2.21551618553357 |
| H | -1.20984530213564 | -4.28146183930034 | 0.93265411596221 |
| O | 1.20282455142790 | -4.24175193884034 | -0.63023744920814 |
| H | 0.71846175869687 | -4.10177731564556 | 0.20208084366272 |
| H | 0.61990675491934 | -4.77167726205029 | -1.18155539008296 |