

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Wallace D. Derricotte

Date

Development and Applications of Orthogonality Constrained Density
Functional Theory for the Accurate Simulation of X-Ray Absorption
Spectroscopy

By

Wallace D. Derricotte
Doctor of Philosophy
Chemistry

Francesco A. Evangelista, Ph.D
Advisor

Dr. Joel Bowman, Ph.D
Committee Member

Dr. Susanna Widicus-Weaver, Ph.D
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D
Dean of the James T. Laney School of Graduate Studies

Date

Development and Applications of Orthogonality Constrained Density
Functional Theory for the Accurate Simulation of X-Ray Absorption
Spectroscopy

By

Wallace D. Derricotte
B.S., Morehouse College, 2013

Advisor: Francesco A. Evangelista, Ph.D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2017

Abstract

Development and Applications of Orthogonality Constrained Density Functional Theory for the Accurate Simulation of X-Ray Absorption Spectroscopy

By Wallace D. Derricotte

The aim of this dissertation is to address the theoretical challenges of calculating core-excited states within the framework of orthogonality constrained density functional theory (OCDFT). OCDFT is a well-established variational, time independent formulation of DFT for the computation of electronic excited states. In this work, the theory is first extended to compute core-excited states and generalized to calculate multiple excited state solutions. An initial benchmark is performed on a set of 40 unique core-excitations, highlighting that OCDFT excitation energies have a mean absolute error of 1.0 eV. Next, a novel implementation of the spin-free exact-two-component (X2C) one-electron treatment of scalar relativistic effects is presented and combined with OCDFT in an effort to calculate core excited states of transition metal complexes. The X2C-OCDFT spectra of three organotitanium complexes (TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2) are shown to be in good agreement with experimental results and show a maximum absolute error of 5-6 eV. Next the issue of assigning core excited states is addressed by introducing an automated approach to analyzing the excited state MO by quantifying its local contributions using a unique orbital basis known as localized intrinsic valence virtual orbitals (LIVVOs). The utility of this approach is highlighted by studying sulfur core excitations in ethanethiol and benzenethiol, as well as the hydrogen bonding in the water dimer. Finally, an approach to selectively target specific core excited states in OCDFT based on atomic orbital subspace projection is presented in an effort to target core excited states of chemisorbed organic molecules. The core excitation spectrum of pyrazine chemisorbed on Si(100) is calculated using OCDFT and further characterized using the LIVVO approach.

Development and Applications of Orthogonality Constrained Density
Functional Theory for the Accurate Simulation of X-Ray Absorption
Spectroscopy

By

Wallace D. Derricotte
B.S., Morehouse College, 2013

Advisor: Francesco A. Evangelista, Ph.D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2017

Acknowledgement

I would like to dedicate this dissertation to my late Aunt Cynthia Stanley who nicknamed me “Doc” since I was a kid. She knew this day would come, I know she is proud. Rest in Peace.

Dedication

After years of intense research, coursework, teaching, training, and learning today is the day. This final note of thanks is the finishing touch, the cherry on top of the cake that lies beneath. As I sit here typing this it is actually hard to believe that a little black kid from East Point, Georgia has grown into his own and will hold the highest degree that the hollow halls of academia has to offer. I am so proud of this accomplishment, but it is not mine alone, I want to take this moment to reflect on all of the people who helped me get this far in my journey.

First off I would like to thank my wife Alencia, she has been a rock for me throughout this process. She has motivated me and guided me through this process in more ways than she knows, I can honestly say that without her love and support I would not be where I am today. I would like to thank my mother and father Glenda and Charles Derricotte who also been an excellent support system and always encouraged me to pursue my dreams. My brother and sister-in-law Charles and Quenetta Derricotte have also been there for me when I needed them. My wonderful niece and nephews Alaria, Micheal, Amir, C.J., and Daniel are constant reminders of how fun it was to be a carefree kid, seeing them grow up during this time has been a great joy and motivation for me.

I want to thank my advisor Francesco Evangelista for accepting me in his group. Francesco and I both arrived at Emory during the summer of 2013, albeit at two different points in our careers, we were both looking to get a great start at Emory. When I first began I knew little to nothing about computational chemistry, but Francesco was very helpful and taught me everything I needed to know to succeed. I am forever grateful for his patience and guidance. I also want to thank all of the Evangelista group members who were here during my tenure with the group Dr. Prakash Verma, Dr. Chenyang Li, Jeffery Schriber, Tianyuan (Sam) Zhang, and Kevin Hannon. I would also like to thank the greater Southeastern Theoretical Chemistry (SETCA) community that has embraced me and been of great help especially Trent Parker, Brandon Bakr, Dr. Daniel Smith, Prof. David Sherrill, Prof. Ryan Fortenberry, and countless others. This community has felt like home to me and hope to be a part of it for years to come.

Table of Contents

1	Introduction and Literature Review	1
1.1	Introduction	1
1.2	Photochemistry and Core Electron Excitations	3
1.3	X-Ray Absorption Spectroscopy	5
1.4	Theoretical Approaches for Calculation of Core Excited States	10
1.4.1	Hartree-Fock Static Exchange	11
1.4.2	Linear Response Time-Dependent Density Functional Theory Approaches	13
1.4.3	Coupled Cluster Approaches	17
1.5	Orthogonality Constrained Density Functional Theory	20
1.5.1	Original Formulation via Constrained Variational Minimization	21
1.5.2	Attractive Features of OCDFT for Core Excitations	22
1.6	Prospectus	26
2	Simulation of X-Ray Absorption Spectra with Orthogonality Constrained Density Functional Theory	33
2.1	Introduction	34
2.2	Theory	37
2.3	Computational Details	43
2.4	Results and Discussion	45
2.4.1	Calibration of OCDFT core-excitation energies	45
2.4.2	Application to Nucleobases: Thymine and Adenine Near-Edge Spectra	51
2.5	Conclusions	60
2.6	Acknowledgements	63
3	Predicting Near Edge X-ray Absorption Spectra with the Spin-Free Exact-Two-Component Hamiltonian and Orthogonality Constrained Density Functional Theory	73
3.1	Introduction	74
3.2	Theory	77
3.2.1	One-Electron Spin-Free X2C	77
3.2.2	Orthogonality Constrained DFT	79
3.2.3	Comparison of OCDFT and TDDFT for Core Excitations	81

3.3	Computational Details	87
3.4	Results and Discussion	88
3.4.1	Calibration of X2C-OCDFT Core Excitation Energies	88
3.4.2	Ti K-Edge NEXAS of Organotitanium Complexes	92
3.5	Discussion and Conclusions	99
4	Localized Intrinsic Valence Virtual Orbitals as a Tool for the Automated Classification of Core Excited States	108
4.1	Introduction	109
4.2	Theory	112
4.2.1	Construction of Localized Intrinsic Valence Virtual Orbitals	113
4.2.2	Determination of the character of the IAOs and LIVVOs	115
4.2.3	Analysis of OCDFT Particle Orbitals Using LIVVOs	117
4.3	Computational Details	118
4.4	Results	119
4.4.1	Analysis of substituent effects in the spectra of thiols	119
4.4.2	Signatures of hydrogen bonding in the NEXAFS spectrum of the water dimer	124
4.4.3	Basis Set Dependence of LIVVO Analysis	131
4.5	Conclusions	133
5	A Maximum Subspace Occupation Approach for the Study of the NEXAFS Spectra of Chemisorbed Or- ganic Molecules Using Orthogonality Constrained Den- sity Functional Theory: Pyrazine on Si(100) a Case Study	141
5.1	Introduction	142
5.2	Method	148
5.2.1	Assigning Transitions Based on Localized Intrinsic Valence Vir- tual Orbitals	151
5.3	Computational Details	152
5.4	Results and discussion	154
5.5	Conclusions and Future Work	161
6	Concluding Remarks and Outlook	169

List of Figures

1.1	Sketch of the difference between UV and X-ray photo-absorption. The energies are not true to scale. The valence excitation and valence ionization processes are shown in blue and red respectively. The core excitation and core ionization processes are shown in green and purple respectively. An example is shown of the first valence and core excited state in ammonia (NH_3), the orbitals shown are hole and particle from orthogonality constrained density functional theory excited state computations at the B3LYP/3-21G level of theory.	4
1.2	a) Schematic highlighting the X-ray absorption process, specifically shown here is an excitation of an electron from the first orbital shell (K-shell) of an arbitrary molecule. b) Drawing of the general shape of a K-edge X-ray absorption spectrum. The unique regions of the spectrum are marked and grouped with their corresponding spectroscopic method, either near-edge (NE) or extended-edge (E) x-ray absorption fine spectroscopy (XAFS).	7
1.3	First O 1s core excitation in H_2O calculated using orthogonality constrained density functional theory at the B3LYP/3-21G level of theory. Numbers reported are the total energy of the ground state and respective core excited state calculated with varying degrees of orbital relaxation included.	24
2.1	Example of a X-ray photoabsorption spectrum (XAS). The near edge, located in the low energy region, consists of excitations of core electrons to valence orbitals. These transitions are sensitive to the chemical environment surrounding the excited atom. The high energy region of the spectrum results from excitations of core electrons to the continuum.	34
2.2	Illustration of the projection operators used in OCDFT. Notice, that a consequence of Eqs. (2.4)–(2.5) is that the hole $[\phi_h^{(1)}]$ and particle $[\phi_p^{(1)}]$ orbitals are contained in the spaces spanned by $\hat{P}^{(0)}$ and $\hat{Q}^{(0)}$	39
2.3	Constrained multiple hole/particle (CMHP) algorithm illustrated in the case of two hole and three particle orbitals.	40

2.4	Histogram showing the distribution of the error in the computed core excited states. All calculations were done using the B3LYP functional and the def2-QZVP basis set. The red filled bar is OCDFT while the empty blue bar is TDDFT. Notice that the two distributions do not overlap.	46
2.5	Scatterplot displaying the excitation energy error as a function of orbital overlap for a benchmarking test set of 40 core excitations from 13 different molecules. Excitation energies were calculated using the B3LYP functional and def2-QZVP basis set. The overlap integrals were computed with numerical grid integration making use of the Gaussian cube files produced by in Psi4 OCDFT calculations. Grids were calculated with a double zeta basis set and 0.1 grid spacing.	47
2.6	Relevant virtual orbitals for thymine numbered in according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is diffused are labeled as D.	53
2.7	Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110.	54
2.8	Relevant virtual orbitals for adenine numbered according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is mostly diffused on the outside of the molecule are labeled as ϕ_D	57
2.9	Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110.	58
3.1	Outline of Orthogonality Constrained DFT: (a) In the Ayers–Levy–Nagy time-independent variational formulation of DFT each electronic is associated with a unique density functional. (b) For each electronic state $\Psi^{(n)}$, an auxiliary system of non-interacting electrons with wave function $\Phi^{(n)}$ is introduced. $\Psi^{(n)}$ and $\Phi^{(n)}$ are assumed to have the same electron density $\rho^{(n)}$. (c) We then impose explicit orthogonality conditions onto the auxiliary wave functions and approximate the unknown exchange-correlation (xc) functional $E_{xc}^{(n)}$ with the ground-state functional. (d) The final OCDFT functional consists in the ground state functional augmented with an orthogonality constraint.	80

3.2	Relative Peak Positions of the C (left) and O (right) K-edge spectra of CO. Spectra are shifted by the amount required to align the first peak to the experimental peak values (287.40 and 534.10 eV for the C and O spectrum respectively). All calculations use the aug-cc-pVQZ basis set.	83
3.3	Error distribution in the computed core excited states for: (a) nonrelativistic OCDFT, (b) OCDFT with the second-order Douglas–Kroll–Hess Hamiltonian, and (c) OCDFT with the exact-2-component Hamiltonian. All calculations use the B3LYP functional and a fully uncontracted aug-cc-pVDZ basis set. Red bars indicate core excitations from 1s orbitals of second-row elements.	89
3.4	(a) Near-edge X-ray absorption spectrum of TiCl ₄ , TiCpCl ₃ , and TiCp ₂ Cl ₂ computed with X2C-OCDFT at the B3LYP/un-cc-pVDZ level of theory. (b)–(d) Comparison of the onset of the pre-edge experimental spectra (dotted line) with the X2C-OCDFT spectrum shifted by +4.7 eV to match the experimental pre-edge peak of TiCl ₄ (solid line). Experimental curves taken from ref 121. Ti K-edge for each molecule was simulated using 25 unique core-excited state (see supplementary material to see all excited states).	95
3.5	(a) Particle orbitals for the core excited states that contribute to the pre-edge features of TiCl ₄ , TiCpCl ₃ , and TiCp ₂ Cl ₂ and (b) corresponding spectral features.	96
4.1	Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOs.	113
4.2	Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software. ⁹⁴	121
4.3	Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) ethanethiol (C ₂ H ₅ SH) and (B) benzenethiol (C ₆ H ₅ SH). Percentages shown represent the overlap of each particle orbital with the LIVVOs. Particle orbitals are numbered according to the calculated transitions reported in Table 4.2 while atom numberings correspond to those shown in Figure 4.2.	122

4.4	Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red circle. Structures were plotted using the CYLview software. ⁹⁴	126
4.5	Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3 while atom numberings correspond to those shown in Figure 4.4.	130
5.1	Molecular structure of pyrazine ($C_4H_4N_2$) with conventional numbering scheme.	142
5.2	Schematic of three unique absorption modes of pyrazine ($C_4H_4N_2$) on Si(100), a) one with a single N atom attached at the Si surface dimer, and two structures that form a 1,4-N-N-dihydropyrazine (DHP)-like structure through forming two covalent Si-N bonds on a b) single dimer or c) across two dimer rows.	143
5.3	Example application of the maximum subspace occupation method within OCDFT to target the C 1s core excitations in CO at the B3LYP/3-21G level of theory: a) the 7 occupied orbitals of CO in C_{2v} symmetry, with the corresponding subspace occupation number (Ω_i) for a subspace composed of the C 1s atomic orbital. b) These projections are used to build the subset of occupied orbitals, in this example a hole threshold of 0.2 is used, and thus orbital $2A_1$ is chosen as the target hole orbital. This allows for direct access to the relevant C K-edge transitions.	150
5.4	Near-edge X-ray absorption spectrum of pyrazine ($C_4H_4N_2$) at the a) C K-edge and b) N K-edge computed with OCDFT (blue line) at the PBE/cc-pVTZ level of theory. Experimental results (black line) are obtained from Ref. 88 and appear here in their original peak position with no applied energy shifts. The pyrazine geometry optimized at the B3LYP/6-31G* level is shown in the inset, the atoms circled (C_1 and N_4) are the core holes considered in further discussion of the results in this section.	154
5.5	Particle orbitals for the core excited states that contribute to the a) carbon and b) nitrogen 1s excitation spectrum. Also shown are c) the LIVVOs that are used to analyze the particle orbitals.	156

5.6	Near-edge X-ray absorption spectrum of pyrazine ($C_4H_4N_2$) chemisorbed on a $Si_{23}H_{24}$ cluster model of Si(100) at the a) C K-edge and b) N K-edge computed with OCDFT at the PBE level. All C, N, and H atoms use the cc-pVTZ basis set while the 6-31G basis set is used for the Si atoms of the cluster model. The inset in the C spectrum magnifies the σ^* manifold of the spectrum. The pyrazine geometry optimized at the B3LYP/6-31G* level is shown in the inset in the N spectrum, the atoms circled (C_8 and N_4) are the core holes considered in further discussion of the results in this section.	158
5.7	LIVVOs used to analyze particle orbitals in pyrazine ($C_4H_4N_2$) chemisorbed on a $Si_{23}H_{24}$ cluster model of Si(100).	159
5.8	Particle orbitals for the core excited states that contribute to the a) carbon and b) nitrogen 1s excitation spectrum of pyrazine ($C_4H_4N_2$) chemisorbed on a $Si_{23}H_{24}$ cluster model of Si(100).	161

List of Tables

1.1	Non-relativistic (NR) and relativistic (R) energies (in eV) of the lowest lying core orbital and lowest unoccupied molecular orbital (LUMO) of water (H ₂ O) and silane (SiH ₄) molecules calculated using spin-unrestricted Hartree–Fock (UHF) with the 3-21G basis set. Scalar relativistic effects are included by use of the 2 nd order Douglas-Kroll-Hess Hamiltonian. ¹⁰³ All calculations were performed using the ORCA software package.	10
2.1	OCDFT core-excitation energies for a benchmark set composed of 13 diatomic molecules. Mean absolute error (in eV) computed using various combinations of basis sets and density functionals. These statistics refer to a subset of the benchmark set comprised of 35 core-excited states.	45
2.2	Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69–77.	49
2.3	Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78–83.	50
2.4	Calculated (B3LYP/def2-TZVP) and experimental thymine oxygen, nitrogen, and carbon core excitation energies (ω_{fi} , in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_h) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_p).	70
2.5	Comparison of the most intense transitions at each edge of the thymine and adenine spectra. OCDFT assignments and energies (in eV) are shown and are compared to the most intense transitions from the theoretical ADC(2) spectra reported in Ref 110.	71

2.6	Calculated (B3LYP/def2-TZVP) and experimental adenine nitrogen and carbon core excitation energies (ω_{fi} , in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_h) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_p).	72
3.1	Singlet Excitation Energies for a Toy Model of Two Electrons in Three Molecular Orbitals.	84
3.2	Excitation energies of the three lowest energy 1s core excitations for each atom in CO calculated with X2C-OCDFT using the B3LYP functional and fully uncontracted cc-pVXZ and aug-cc-pVXZ basis sets ($X = \text{D, T, Q}$). Experimental values are taken from Refs. 107 and 110. Errors greater than 1 eV are indicated with bold font.	88
3.3	OCDFT core excitation energies for molecules containing first-row elements using nonrelativistic (NR) and relativistic (X2C,DKH2, DKH3, DKH4) Hamiltonians. Computations were performed using the uncontracted aug-cc-pVDZ basis with the B3LYP exchange-correlation functional ($E_{\text{xc}}^{\text{B3LYP}}$) and 100% non-local exact exchange (E_{x}^{HF}). OCDFT results are reported here as deviations from the experimental excitation energies (in eV). The mean absolute error (MAE) is also reported for each method. Experimental values are taken from refs 105–114.	90
3.4	OCDFT core excitation energies for molecules containing first-row elements using nonrelativistic (NR) and relativistic (X2C,DKH2, DKH3, DKH4) Hamiltonians. Computations were performed using the uncontracted aug-cc-pVDZ basis with the B3LYP exchange-correlation functional ($E_{\text{xc}}^{\text{B3LYP}}$) and 100% non-local exact exchange (E_{x}^{HF}). OCDFT results are reported here as deviations from the experimental excitation energies (in eV). The mean absolute error (MAE) is also reported for each method. Experimental values are from refs 115–120	91
3.5	Importance of relativistic and correlation effects in core excitations computed with X2C-OCDFT. Relativistic effects are estimated as the difference between NR and X2C OCDFT calculations [$\text{OCDFT}^{\text{NR}}(E_{\text{xc}}^{\text{B3LYP}}) - \text{OCDFT}^{\text{X2C}}(E_{\text{xc}}^{\text{B3LYP}})$], while correlation effects are estimated as the difference between B3LYP and HF OCDFT calculations [$\text{OCDFT}^{\text{X2C}}(E_{\text{xc}}^{\text{B3LYP}}) - \text{OCDFT}^{\text{X2C}}(E_{\text{x}}^{\text{HF}})$]. Molecules with symmetry equivalent nuclei were excluded from this analysis due to inability of the HF functional to obtain a broken-symmetry solution (see text for further details).	93
3.6	TiCl ₄ pre-edge excitation energies computed with TDDFT and OCDFT using nonrelativistic (NR) and relativistic (DKH2,X2C) Hamiltonians. Results are reported as deviations from the experimental peak maximum (in eV, from ref 121). All calculations use the B3LYP functional and the uncontracted cc-pVDZ basis set.	93

3.7	Transition energies (in eV) and relative oscillator strengths (f) for pre-edge transitions of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 computed with X2C-OCDFT at the B3LYP/un-cc-pVDZ level of theory. Oscillator strengths shown are values relative to the most intense Ti K-edge transition of these three compounds (see supplementary material for detail).	95
3.8	Change in the natural orbital population numbers of the Ti p, Ti d, Cl p shells computed for the Ti pre-edge excitations of TiCl_4 , TiCpCl_3 , TiCp_2Cl_2 .	97
3.9	Partial atomic contributions to the transition dipole moment (in atomic units) computed for the pre-edge features of TiCpCl_3 .	99
4.1	Atom-centered gross orbital populations [GP_{A_i}] for the seven IAOs of water. Calculations were performed at the B3LYP/jun-cc-pVTZ level of theory.	116
4.2	Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57.	120
4.3	Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-D.	125
4.4	Basis set dependence of the IAO assignment of LIVVOs and overlap [$\Omega_{\text{pl}}^{(n)}$] of the $\sigma_{\text{S}^3\text{-H}^9}^*$ and $\sigma_{\text{S}^3\text{-C}^2}^*$ LIVVOs with the first two particle orbitals of ethanethiol.	133
5.1	First 3 carbon core excitation energies and oscillator strengths (f_{osc}) in a CCl_4 molecule calculated with and without the maximum subspace occupation (MSO) method in OCDFT. All calculations were performed using the B3LYP functional and aug-cc-pvdz basis set.	151
5.2	Calculated (PBE/cc-pVTZ) core excitation energies, oscillator strengths, and LIVVO assignments for the C and N 1s excitation spectrum of free pyrazine. The LIVVO assignments shown are the two with the highest percentage overlap with the particle orbital. For each calculated transition we report the π^*/σ^* mixing ratio (see text for details) and the spectral feature that this state contributes to (see Figure 5.4).	155
5.3	Calculated (PBE/cc-pVTZ) core excitation energies, oscillator strengths, and LIVVO assignments for the C and N 1s excitation spectrum of pyrazine ($\text{C}_4\text{H}_4\text{N}_2$) chemisorbed on a $\text{Si}_{23}\text{H}_{24}$ cluster model of Si(100). The LIVVO assignments shown are the two with the highest percentage overlap with the particle orbital. For each calculated transition we report the π^*/σ^* mixing ratio (see text for details) and the spectral feature that this state contributes to (see Figure 5.6).	162

Chapter 1

Introduction and Literature Review

“In science one tries to tell people, in such a way as to be understood by everyone, something that no one ever knew before. But in the case of poetry, it’s the exact opposite!”

— Paul A. M. Dirac

1.1 Introduction

Detailed study and characterization of electronic excited states is an active branch of research in the chemical sciences for both experimentalists and theoreticians alike.¹⁻⁴ To obtain a truly robust understanding of the properties and reactivity of molecules it is often necessary to study their behavior beyond the ground electronic state. Scores of spectroscopic techniques are actively pursued to probe excited electronic states,⁵ with accurate theoretical calculations being a crucial component for the full understanding of spectral signals and what they mean with regards to molecular electronic structure.^{6,7} Depending upon the nature of the excitation, its theoretical treatment can be extremely challenging.⁸⁻¹⁰ High-energy excitations of electrons closest to the atomic nuclei (1s,2s,2p,etc.) are an example of a class of excitations that require a nontrivial theoretical treatment.¹¹ These core-level excitations occur when a molecule absorbs an X-ray photon and an electron is excited from a core level to an unoccupied level.¹² Studying the nature of core-excited states and relating their behavior to

the electronic properties of molecular systems encompass a field of study known as X-ray absorption spectroscopy (XAS).¹³ Modern developments in synchrotron radiation^{14,15} have made XAS an accessible and essential tool for evaluating local electronic and geometric structure, with experimental and theoretical developments both being necessary to its usefulness as an analytical technique.^{16,17}

Rigorous interpretation of XAS requires knowledge of excitation energies, transition dipole moments, and orbital character of core-level excited states. Computation of these and other important properties are the focus of numerous approaches in quantum chemistry. However, excitations involving core electrons present a unique set of computational challenges. Orbitals that comprise the core level are energetically well separated from the rest of the electronic structure and strongly contracted toward the nuclei due to attractive Coulombic forces. Thus, generating a core hole results in a decrease in the shielding of the nuclei¹⁸ causing significant rearrangement of the orbitals in the valence level.¹⁹ This rearrangement is known as orbital relaxation, and it contributes to a significant lowering of the total energy of core-excited states.²⁰ An additional challenge for calculating core excited states is the need for the treatment of relativistic effects.^{17,21} Due to their close proximity to the atomic nucleus, core orbitals experience a significant lowering of energy due to relativistic effects while the valence orbitals remain fairly unchanged. This results in an increase in excitation energy with respect to a nonrelativistic solution that is often non-negligible and becomes particularly important for core excited states of heavier nuclei.²² Along with obtaining accurate energies and properties, it is equally important for computational approaches to provide a detailed description of the nature of each core transition.²³⁻²⁵ Thorough interpretation of the X-ray absorption spectrum involves assigning spectral features based on the classification of the final state. Molecular orbital (MO) theory is often useful in this context, however, the orbital character can potentially be difficult to discern from MO plots leading to ambiguous or undefined assignments.²⁶⁻²⁸

More detailed characterizations are desired in order to obtain a better description of spectral contributions.

The aim of this dissertation is to address all of the key theoretical challenges of calculating and classifying core-excited states within the framework of orthogonality constrained density functional theory (OCDFT). This theory is a previously established²⁹ method for studying electronic excited states. In the scope of this dissertation, I have worked toward enhancing the capabilities of OCDFT in order to properly treat core-excitations, and applied the theory to interesting chemical problems within XAS. This introduction will serve as a brief overview of x-ray absorption spectroscopy, core-excitations, and orthogonality constrained density functional theory. I will begin with an overview of the nature of electronic excited states and the unique nature of core-excited states. What follows is a brief introduction of X-ray absorption spectroscopy (XAS) and a selection of the current theoretical methods used for spectral simulation. This section closes with an introduction to OCDFT and specifically what makes it an excellent choice for calculating this class of excitations.

1.2 Photochemistry and Core Electron Excitations

The study of the interaction of atoms and molecules with electromagnetic radiation constitutes a unique branch of chemistry known as photochemistry. Examination of photochemical processes has practical value for a multitude of different fields. For example, in cancer biology studying the interaction of radiation with DNA aides in understanding the photo-initiated process of DNA damage that leads to cancer formation. In organoelectronics, a whole host of efficient organic solar cells and light emitting diodes are being developed by understanding light-absorbing molecules. Even in nature, photochemistry is abundant through many processes, including the formation of Vitamin D with sunlight,³⁰ photosynthesis,³¹ and vision.³²

If we consider a closed-shell molecule, i.e. where all electrons are paired, when

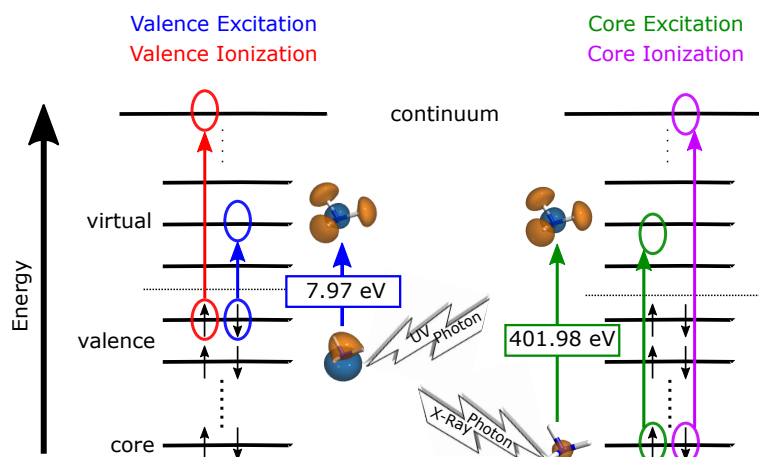


Figure 1.1: Sketch of the difference between UV and X-ray photo-absorption. The energies are not true to scale. The valence excitation and valence ionization processes are shown in blue and red respectively. The core excitation and core ionization processes are shown in green and purple respectively. An example is shown of the first valence and core excited state in ammonia (NH₃), the orbitals shown are hole and particle from orthogonality constrained density functional theory excited state computations at the B3LYP/3-21G level of theory.

this molecule is not under the influence of any external radiation it is in its ground electronic state (S_0). Photons from impending radiation can excite the molecule from the ground state to an electronic excited state, promoting an electron from an occupied molecular orbital (MO) to a higher energy unoccupied MO (often referred to as a “virtual” orbital in quantum chemistry). The excitation energy (ω) of the electronic excited state will depend upon the frequency of the impending radiation. The lifetime of these electronic excited states are extremely short compared to the lifetime of molecular vibrations and thus these states can often be modeled properly within the vertical Franck-Condon approximation^{33–36} where the electronic transition is assumed to occur without significant changes to the positions of the nuclei of the molecular system. The molecule eventually relaxes back down to its ground state level via a subsequent decay process such as fluorescence,^{37,38} phosphorescence,^{39–41} radiationless internal conversion,^{42,43} intersystem crossing,^{44,45} and a myriad of other

processes.^{46–49}

A unique class of electronic photo-excitations involves promotion of a core electron to a valence MO with the corresponding state often referred to as core-excited (core-excited states). The electrons that constitute the “core” of a molecule are those occupying the lowest energy molecular orbitals. These MOs are not actively involved in bonding with the surrounding molecular environment and have orbital character similar to free-atom atomic orbitals (i.e. the 1s core orbital of NH_3 in Figure 1.1). The number of MOs that compose the core is dependent upon the atomic number (Z) of the atom involved. For example, for lighter first-row elements in the periodic table, like oxygen, the 1s orbital is the only core orbital. However for heavier elements, like titanium, the 2s and 2p are also considered in the core level. A substantial amount of energy is required to excite electrons from the core level due to their strong Coulombic attraction to the nucleus. Thus, while ultraviolet (UV) or visible (VIS) photons are sufficient to excite valence electrons, high energy X-ray photons are required in order to excite core electrons. Figure 1.1 shows this process for both the valence and core cases respectively, and details the energy difference between the two in an example case of ammonia (NH_3). The first electronic valence excitation in ammonia has an excitation energy of 7.97 eV, meanwhile the first N core excitation has an excitation energy of 401.98 eV. It is clear that the high energy range, and unique nature of these transitions, warrants an entire field of spectroscopy dedicated to its understanding.

1.3 X-Ray Absorption Spectroscopy

Core-excited states can be probed experimentally via X-ray absorption spectroscopy (XAS), a rapidly growing field that commonly makes use of synchrotron radiation in order to provide a tunable and intense X-ray beam source.^{14,15} Tunability of the X-ray beam source allows it to scan a wide energy range throughout the X-ray region.⁵⁰ Scanning the X-ray beam through the binding energy of a core electron results in an

abrupt increase in the absorption cross-section. This generates the so-called absorption “edge”, with each unique edge region corresponding to a different core electron binding energy.¹³ The different edge regions are named using standard X-ray notation which is based on the principal quantum number of the excited electron; K for $n=1$, L for $n=2$, and M for $n=3$ (see Figure 1.2a).¹² Each edge is well separated from the rest of the absorption spectrum due to the dependence of the binding energy of core electrons on the charge of the nucleus. Since the edge position is dependent upon the nuclear charge it has an inherent dependence on the atomic number (Z). This is why XAS is often referred to as an “element specific” technique. For example, a typical K-edge resulting from oxygen core excitations is located at approximately 530 eV,^{51,52} while the K-edge for nitrogen is located at approximately 400 eV.^{53,54} The existence of an absorption edge by itself has little relevance outside of simple elemental identification.⁵⁵ Fortunately the edge is not simply a sharp absorption increase, it has an abundance of structure (often referred to as “fine structure”) throughout the different regions of the edge. Figure 1.2b shows a schematic drawing of a typical K-edge with the different regions of the fine structure highlighted. The first region of fine structure just before the edge is known as the pre-edge region. Distinct features only arise in this region in the case of transition metal complexes, where this region is populated by core transitions to unoccupied d-orbitals.^{56–59} The rising-edge region is composed of all other discrete core transitions of the system and has a distinct shape, that is typically dominated by intense $1s \rightarrow \pi^*$ transitions accompanied by weaker $1s \rightarrow \sigma^*$ transitions. The study of the fine structure that composes these two edge regions are the focus of near-edge X-ray absorption fine structure spectroscopy (NEXAFS).¹³ The last region of edge structure, the extended-edge region, is composed of signals past the ionization threshold of the core orbital in which core electrons are promoted to continuum (see Figure 1.1). At this point the excited photoelectron has a high amount of energy, enough for its wavelength to be on the scale of inter-

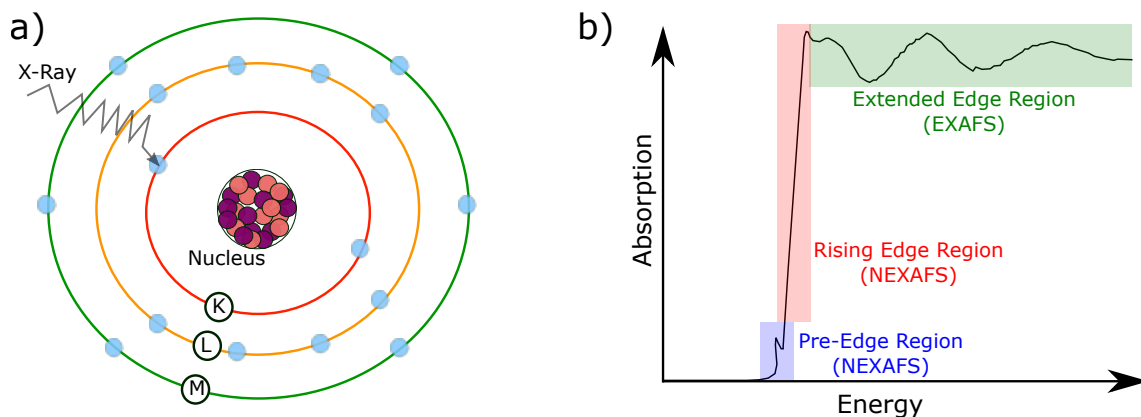


Figure 1.2: a) Schematic highlighting the X-ray absorption process, specifically shown here is an excitation of an electron from the first orbital shell (K-shell) of an arbitrary molecule. b) Drawing of the general shape of a K-edge X-ray absorption spectrum. The unique regions of the spectrum are marked and grouped with their corresponding spectroscopic method, either near-edge (NE) or extended-edge (E) x-ray absorption fine spectroscopy (XAFS).

atomic distances. The resulting spectrum is a characteristic pattern of modulating local maxima and minima that correspond to constructive and destructive interference between outgoing and backscattered photoelectron waves. As a result of this atomic-level specificity of the signal, this region is sensitive to the radial distribution of electron density around the atom. The study of this signal, and its use to quantify geometric information such as bond lengths/coordination numbers is known as extended-edge X-ray absorption fine structure spectroscopy (EXAFS).^{60–63}

With regard to the electronic structure of molecules, the near-edge region contains very useful information.^{64,65} NEXAFS has been applied to a wide range of molecular systems including biological systems,^{66–68} small gas phase molecules,^{69–71} organic thin-films,^{72–74} and semiconducting materials.⁷⁵ For example, information regarding the orientation of organic molecules chemisorbed to surfaces can be obtained with NEXAFS.^{76–79} One study investigated the adsorption structure of pyrazine ($C_4H_4N_2$) on a reconstructed Si(100) surface.⁸⁰ Through investigating changes in the polarization dependent NEXAFS spectrum, the geometry of the molecule relative to the surface

normal could be deduced. NEXAFS can even be useful in explaining reaction mechanisms by using time-resolved X-ray spectroscopy.⁸¹ A recent work by Attar et al.⁸² made use of time-resolved femtosecond NEXAFS in order to study the electrocyclic ring-opening reaction of 1,3-cyclohexadiene. Tracking the changes in the NEXAFS spectrum over time revealed strong mixing amongst the frontier orbitals and provided justification for the long postulated Woodward-Hoffman rules for pericyclic reaction mechanisms.⁸³

Detailed interpretation of experimental X-ray absorption spectra requires knowledge of the relevant core excitation energies and excited state properties. Accurate quantum chemistry methods can often calculate electronic excited states and simulate the absorption spectrum of molecules. Along with interpretation of the properties and character of the transition, quantum chemistry calculations can yield plenty of useful information regarding the nature of electronic excited states. Many modern computational techniques have the capability of producing quantitatively accurate results regarding the excited states of systems of moderate size (less than 20 atoms).⁸⁴⁻⁸⁷ However, for the excited states of significantly larger systems, approximations must be made that represent a concession of accuracy for the sake of computational efficiency.^{88,89} Utilizing quantum chemistry methods for the study of core-excited states encounters quite a few unique challenges. The most practical challenge they face from a purely computational perspective involves the eigenvalue solvers employed by many methods. Iterative eigenvalue solvers are routinely employed in many quantum chemistry methods.⁹⁰⁻⁹³ These approaches typically yield the lowest eigenvalue solutions and thus, tend to be efficient approaches for solving valence excited states. Core-excited states, however, are much higher in energy. In order to obtain solutions from the core region of the excitation spectrum, all underlying valence excited states must be calculated.⁹⁴ This is computationally inefficient in all cases and impossible for larger systems. Multiple strategies exist to overcome this algorithmic challenge, they

typically revolve around target specific energy solutions⁹⁵⁻⁹⁷ or formally decoupling the space of core excitations from the space of valence excitations.^{98,99} Both routes are viable options for overcoming this difficulty and the details of these approaches will be reviewed throughout the following section. Other challenges faced by quantum chemistry methods when calculating core-excited states are more physical in nature and have to do with the unique physics present in the core excitation mechanism. As stated previously, core electrons are tightly bound due to their close proximity to the nuclei and generation of a core hole causes significant orbital relaxation and a subsequent lowering of the total energy. Although orbital relaxation technically occurs in every excited state, the effect is less pronounced in valence excitations and is often times negligible in these calculations. To highlight this, a study by Triguero et al.¹⁰⁰ investigated the sources of error in Koopmans approximation¹⁰¹ of an ionization potential, I . This error can be divided into three contributions: $R1$ (< 0) due to orbital relaxation, $R2$ (> 0) due to correlation, and $R3$ (< 0) due to the relaxation induced change in correlation. For valence ionizations, the relationship between the three sources of error is $-R1 \approx R2 \gg -R3$. However, for core ionizations this relationship becomes $-R1 \gg R2 \approx -R3$ with the error due to orbital relaxation becoming the dominant term. For example, Plakhutin et al.¹⁰² showed that for core ionization of a nitrogen atom, $R1$ is on the order of 12-15 eV.¹⁰² Clearly this energy contribution is non-negligible for core-excited states and any quantum chemistry method striving for accurate energy calculations must consider this effect. Another consideration is the treatment of relativistic effects. Relativistic contraction of the core orbitals toward the nuclei results in a lowering of the core orbital energy while the valence orbital energies remain relatively unaffected. This results in an increase in excitation energy with respect to a nonrelativistic solution that is often non-negligible and becomes particularly important for core-excited states of heavier nuclei. Table 1.1 shows orbital energies of the lowest lying core orbital and the lowest unoccupied molecular

Table 1.1: Non-relativistic (NR) and relativistic (R) energies (in eV) of the lowest lying core orbital and lowest unoccupied molecular orbital (LUMO) of water (H₂O) and silane (SiH₄) molecules calculated using spin-unrestricted Hartree–Fock (UHF) with the 3-21G basis set. Scalar relativistic effects are included by use of the 2nd order Douglas-Kroll-Hess Hamiltonian.¹⁰³ All calculations were performed using the ORCA software package.

	H ₂ O		SiH ₄	
	O 1s	LUMO	Si 1s	LUMO
NR-UHF	-555.88	7.17	-1858.20	5.02
R-UHF	-556.34	7.16	-1863.04	5.03

orbital (LUMO) for water (H₂O) and silane (SiH₄) molecules in a non-relativistic SCF calculation and a relativistic computation where scalar relativistic effects are taken into account using the Douglas–Kroll Hess Hamiltonian.¹⁰³ The core orbital energy in H₂O lowers by 0.46 eV while the LUMO experiences negligible energy lowering of 0.01 eV. This effect becomes more pronounced with heavier elements, as the core orbital of silane experiences a relativistic contraction of 4.84 eV while its LUMO raises in energy by only 0.01 eV. In the following section I will review the current theoretical approaches that have been applied to core-excited states and address how they approach the treatment and/or approximation the necessary components.

1.4 Theoretical Approaches for Calculation of Core Excited States

The advent of synchrotron radiation made XAS experiments more feasible, resulting in increased interest in the technique. This has led to numerous computational approaches for calculating the properties of core-excited states. As discussed in detail in the previous section, computational approaches encounter a wealth of computational challenges when calculating core-excited states, including the algorithmic challenge of targeting the core region of the excitation spectrum and treatment of orbital relaxation and relativistic effects. This section will focus on reviewing theo-

retical approaches for the treatment of core-excited states within the framework of Hartree–Fock static exchange (HF-STEX),^{104–107} time-dependent density functional theory (TDDFT),^{94,97,108} and coupled cluster theory (CC).^{109–113} For information regarding other noteworthy theoretical approaches including algebraic diagrammatic construction (ADC),^{99,114,115} restricted active space self-consistent-field approaches (RASSCF),¹¹⁶ configuration interaction,^{117–120} real-time time-dependent DFT (RT-TDDFT),¹²¹ and/or approaches using the Bethe-Salpeter equation,^{122,123} I refer the reader to the existing literature on these topics. The focus here will be on treatment core-excitations relevant to K-edge applications. It is worth noting that there are numerous approaches for the L-edge that require a more detailed theoretical treatment due to spin-orbit coupling. Detailed treatment of this topic is outside of the scope of this dissertation, and I refer the reader to the existing literature on the topic.^{124–126}

1.4.1 Hartree-Fock Static Exchange

HF-STEX (also known as improved virtual orbital Hartree–Fock) is a state specific Hartree–Fock approach in which a simplified configuration interaction singles expansion is performed to obtain excited states from a reference determinant.^{104–107} The reference determinant is constructed by performing an open-shell HF calculation in which an electron is removed from the core region of an N-electron system. This direct inclusion of relaxation effects is one of the strongest advantages of HF-STEX based methods and is the starting approximation for practically every state specific/variational approach for core-excited state. The validity of this approximation is supported by the large energy separation between the core and valence orbitals as well as the spatial separation of core orbitals on different atomic centers within a molecule. Starting from a closed shell reference state (Ψ_0), let us consider an excitation from core orbital j to valence orbital b (Ψ_{jb}). The STEX eigenvalue problem

can be defined as

$$F^j \phi_b = \epsilon_{jb} \phi_b, \quad (1.1)$$

where F^j is the HF-STEX Fock operator

$$F^j = h + \sum_{i \neq j} (2J_i - K_i) + J_j - K_j. \quad (1.2)$$

It should be noted that the excited states obtained from the open-shell HF reference are not orthogonal to the HF ground state and thus overlaps between the states must be accounted for. After accounting for non-orthogonality, the HF-STEX eigenvalue equation can be solved and the excitation energies (ω_{jb}) can be obtained by adding the core ionization potential to the eigenvalue obtained in Equation 1.1. The oscillator strengths are evaluated in terms of the excitation energy and the transition dipole moment as

$$f_{jb} = \frac{2}{3} \omega_{jb} |\langle \Psi_0 | \vec{r} | \Psi_{jb} \rangle|^2, \quad (1.3)$$

With regard to the performance of HF-STEX, it produces XAS spectra that are in acceptable quantitative agreement with experimental results. HF-STEX oscillator strengths also provide a good qualitative comparison with experimental line intensities.

Excited state solutions within HF-STEX can be best understood as an approximate solution to the random phase approximation (RPA). Lets consider a general RPA excitation operator \hat{T}_I for an excited state I within the Tamm-Dancoff approximation (TDA):¹²⁷

$$\hat{T}_I = \sum_{i,a} (\mathbf{X}_{ia,I} \hat{a}_a^\dagger \hat{a}_i), \quad (1.4)$$

where \mathbf{X} is the eigenvector of the RPA equation within the TDA, while i and a are indices that run over the ground state occupied and virtual orbitals respectively. This type of excitation operator constitutes what is known as a “multi-channel” approach, essentially a linear combination of single excitations that accounts for the interaction of different excitation channels. For HF-STEX, this excitation operator is limited

to a restricted sum over single excitations from a given orbital j , also known as a “single-channel” approach

$$\hat{T}_I = \sum_a (\mathbf{X}_{ja,I} \hat{a}_a^\dagger) \hat{a}_j, \quad (1.5)$$

It should be noted that this would be an extremely poor approximation for valence-excited states as multiple excitation channels tend to be clustered within a fairly small energy range (potentially less than 1 eV). However, due to the large energy separation of core orbitals, it is possible to obtain an accurate simulation of the XAS spectrum within a single channel approximation.

The attractive features of this theory are however hampered by two main challenges. The first is the difficulty of converging the open-shell SCF reference state necessary to build the STEX Hamiltonian. It is notoriously difficult to converge on the desired core hole state. However, there are solutions to this problem that can aide convergence of the core hole state, most notably the maximum overlap method.¹²⁸ The second challenge with HF-STEX lies in the non-orthogonality of the excited states within the core-hole potential which requires explicit wavefunction overlaps to be accounted for.

1.4.2 Linear Response Time-Dependent Density Functional Theory Approaches

With regard to computational cost, linear response time-dependent density functional theory (TDDFT) remains one of the most attractive options for calculating electronic excited states in a wide variety of systems. Within TDDFT, excitation energies (ω) are calculated via solutions to the following non-Hermitian eigenvalue equation:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\dagger & \mathbf{A}^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ -\mathbf{Y} \end{pmatrix}, \quad (1.6)$$

where the matrices \mathbf{A} and \mathbf{B} are given by

$$A_{ai\sigma,bj\tau} = \delta_{ij}\delta_{ab}\delta_{\sigma\tau}(\epsilon_{a\sigma} - \epsilon_{i\tau}) + (ia\sigma|jb\tau) + (ia\sigma|f_{XC}|jb\tau), \quad (1.7)$$

$$B_{ai\sigma,jb\tau} = (ia\sigma|jb\tau) + (ia\sigma|f_{XC}|jb\tau), \quad (1.8)$$

for occupied orbitals i, j, \dots , virtual orbitals a, b, \dots , spin indices σ and τ , and energy eigenvalues ϵ . The second term in Equation 1.7 is a two-electron integral and can be expressed explicitly in terms of a set of Kohn–Sham orbitals (ϕ_i) as:

$$(ia\sigma|jb\tau) = \int \int \phi_{i\sigma}^*(\mathbf{r}_1)\phi_{a\sigma}^*(\mathbf{r}_1)\frac{1}{r_{12}}\phi_{j\tau}^*(\mathbf{r}_2)\phi_{b\tau}^*(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2, \quad (1.9)$$

while the third term is what is known as the exchange correlation kernel which can be expressed as the functional derivative

$$(ia\sigma|f_{XC}|jb\tau) = \int \phi_{i\sigma}^*(\mathbf{r}_1)\phi_{a\sigma}^*(\mathbf{r}_1)\frac{\delta^2 E_{XC}}{\delta\rho_\sigma(\mathbf{r}_1)\delta\rho_\tau(\mathbf{r}_2)}\phi_{j\tau}^*(\mathbf{r}_2)\phi_{b\tau}^*(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2, \quad (1.10)$$

where E_{XC} is the exchange correlation functional. Within the Tamm-Dancoff approximation (TDA),¹²⁹ Equation 1.6 reduces to a Hermitian eigenvalue equation $\mathbf{A}\mathbf{X} = \omega\mathbf{X}$. Eliminating the off diagonal elements \mathbf{B} greatly reduces the computational cost of solving the equation and makes the theory formally similar to configuration interaction singles. Many modern implementations of TDDFT make use of some form of the Davidson algorithm which iteratively diagonalizes a subspace of the Hamiltonian yielding its lowest (highest) eigenvalues in a bottom-up (top-down) fashion. This is a reasonable and efficient choice when the excitations of interest are valence excitations. However, excitations involving the lower energy core orbitals introduce a significant computational challenge since there is typically a plethora of valence excitations below the target core excitations, making the algorithm prohibitively expensive. Overcoming this issue is the primary motivation of methods developed to calculate core excitations within the TDDFT.

Restricted Excitation Window TDDFT

Originally proposed by Stener et al.,¹³⁰ restricted excitation window TDDFT (REW-TDDFT) is a modification to the standard TDDFT algorithm that only allows electrons to be excited from a subset of core orbitals.^{94,108} The REW approximation is

physically motivated by the large energy separation of core levels of different atomic species, and the large spatial separation of core orbitals on different atoms. REW is normally done in one of two ways: (1) Core orbitals of interest are selected directly via orbital localization schemes for equivalent atoms (i.e. selecting all MOs that have primarily oxygen 1s character) or (2) Use an orbital energy (energy difference) cutoff in order to filter out any orbitals (orbital pairs) that lie above (below) the user provided energy threshold. The latter option is by far the most popular and has become commonplace for TDDFT implementations in modern quantum chemistry packages.

Although REW-TDDFT allows for the computation of core excited states, it naturally inherits the issues of standard TDDFT calculations, namely the large dependence on the choice of exchange-correlation functional. This has led to development of a new class of functionals for use with REW-TDDFT known as short range corrected (SRC) functionals, which are modified versions of long range corrected (LRC) functionals used for charge transfer excitations in TDDFT. The LRC exchange correlation functional can be expressed as

$$E_{XC}^{LRC} = E_X^{SR-DFT} + E_X^{LR-HF} + E_C^{DFT}, \quad (1.11)$$

combining a short-range DFT exchange (E_X^{SR-DFT}), with long-range Hartree–Fock exchange (E_X^{LR-HF}), and DFT correlation energy (E_C^{DFT}). SRC functionals modify this general form to include a certain portion of Hartree–Fock exchange in the short range as well, where the SRC exchange-correlation energy (E_{XC}^{SRC}) can be expressed as

$$E_{XC}^{SRC} = c_X E_X^{SR-HF} + E_X^{SR-DFT} + E_X^{LR-HF} + E_C^{DFT}, \quad (1.12)$$

where c_X is a parameter that can be optimized. The major drawback with this method is that the amount of HF exchange in the functional must be re-optimized for any given system, which can be cumbersome and limits its reliability in application to novel problems.

Energy Specific TDDFT

Starting from the non-Hermitian eigenvalue problem shown in Equation 1.6, this form is often placed into the following equivalent Hermitian eigenvalue equation

$$(\mathbf{A} - \mathbf{B})^{1/2}(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B})^{1/2}\mathbf{T} = \omega^2\mathbf{T}, \quad (1.13)$$

where $\mathbf{T} = (\mathbf{A} - \mathbf{B})^{-1/2}(\mathbf{X} + \mathbf{Y})$. Energy Specific TDDFT (ES-TDDFT)⁹⁷ aims to solve this Hermitian eigenvalue problem utilizing a “growing window” algorithm that starts from a set of trial vectors \mathbf{C} associated with orbital pairs whose orbital energy difference are above a given threshold. A subspace $\tilde{\mathbf{M}}$ is then created using the trial vectors,

$$\tilde{\mathbf{M}}^+ = \mathbf{C}^T(\mathbf{A} + \mathbf{B})\mathbf{C} \quad (1.14)$$

$$\tilde{\mathbf{M}}^- = \mathbf{C}^T(\mathbf{A} - \mathbf{B})\mathbf{C} \quad (1.15)$$

$$\tilde{\mathbf{M}} = (\tilde{\mathbf{M}}^-)^{1/2}(\tilde{\mathbf{M}}^+)(\tilde{\mathbf{M}}^-)^{1/2}, \quad (1.16)$$

Direct diagonalization is performed only in the subspace $\tilde{\mathbf{M}}$, which generates a new set of eigenvalues $\tilde{\omega}$ and eigenvectors \tilde{T} . Only those eigenvectors associated with eigenvalues above the user defined threshold are kept. These eigenvectors are then projected back onto the full MO space in order to calculate a residual, any new eigenvectors with significant amplitude are then added as a trial vector for the next iteration (hence the term “growing window”). A new subspace is constructed from the expanded trial space and the process continues self-consistently until the residual is below a user-defined threshold.

Since this method checks the residual of the subspace against the full MO space, the final results are true solutions of the TDDFT equations. Similar to REW-TDDFT, however, it inherits the functional dependence of TDDFT for the calculation of core excited states.

1.4.3 Coupled Cluster Approaches

The coupled-cluster wavefunction (Ψ_{CC}) can be defined as an exponential expansion of a closed shell HF reference (Ψ_{HF})

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}}|\Psi_{\text{HF}}\rangle, \quad (1.17)$$

where \hat{T} is known as the cluster operator can be expressed as a linear combination of single, double, triple, etc. excitations, up to N-fold excitations for a general N-electron system

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_N. \quad (1.18)$$

Each individual excitation operator is composed of cluster amplitudes (t_μ) and excitation operators ($\hat{\tau}_\mu$) that promote electrons from occupied orbitals (i,j,k,...) to virtual orbitals (a,b,c,...)

$$\hat{T} = \sum_{\mu} t_{\mu} \hat{\tau}_{\mu} = \sum_{i>j>k>\dots}^{\text{occ}} \sum_{a>b>c>\dots}^{\text{virt}} t_{ijk\dots}^{\hat{a}bc\dots} \hat{\tau}_{ijk\dots}^{\hat{a}bc\dots}. \quad (1.19)$$

The total energy (E) and amplitudes (Ω_μ) of the ground state can then be determined by projections onto the reference state and onto a manifold of excitations out of the reference state

$$E = \langle \Psi_{\text{HF}} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_{\text{HF}} \rangle, \quad (1.20)$$

$$\Omega_\mu = \langle \Psi_\mu | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_{\text{HF}} \rangle = 0. \quad (1.21)$$

The truncation of the cluster operator \hat{T} gives rise to the different “levels” of coupled cluster theory:

$$\text{CCS} : \hat{T}_1 \quad (1.22)$$

$$\text{CCSD} : \hat{T}_1 + \hat{T}_2 \quad (1.23)$$

$$\text{CCSDT} : \hat{T}_1 + \hat{T}_2 + \hat{T}_3 \quad (1.24)$$

$$\text{CCSDTQ} : \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4 \quad (1.25)$$

⋮

Within coupled-cluster theory, excited states can be obtained via a linear-response approach as well as a state-specific approach. Both approaches have challenges that must be overcome in order to calculate core-excited states. These methods for calculating core-excited states will be reviewed here.

Linear-Response Coupled Cluster Theory

Within coupled-cluster linear response theory (CC-LR), excitation energies (ω_k) are obtained by solving an asymmetric eigenvalue equation to obtain right (\mathbf{R}_k) and left (\mathbf{L}_k) eigenvectors, commonly referred to as excitation and de-excitation operators respectively

$$\mathbf{A}\mathbf{R}_k = \omega_k\mathbf{R}_k, \quad (1.26)$$

$$\mathbf{L}_k\mathbf{A} = \omega_k\mathbf{L}_k, \quad (1.27)$$

where the left and right eigenvalues are orthogonal ($\mathbf{L}_i\mathbf{R}_k = \delta_{ik}$), and \mathbf{A} is a Jacobian matrix that is defined as the gradient of the cluster amplitudes

$$A_{\mu\nu} = \frac{\partial\Omega_\mu}{\partial t_\nu}, \quad (1.28)$$

Many general schemes to iteratively solve Equations 1.26 and 1.27 have been introduced for excited state calculations. These methods typically rely on using HF orbital energy differences to select guess unit vectors for specific occupied to virtual excitations. As with other similar approaches, the iterative solver will converge onto the lowest energy eigenvalues and eigenvectors even if the starting vectors are chosen for higher energy excitations. To overcome this difficulty, Coriani and Koch introduced a novel approach to iteratively solving Equations 1.26 and 1.27 using an asymmetric Lanczos algorithm.^{111,112} This approach employs a tridiagonal representation \mathbf{T} of the Jacobian matrix \mathbf{A} .

$$\mathbf{T} = \mathbf{P}^T\mathbf{A}\mathbf{Q}, \quad (1.29)$$

where $\mathbf{P}^T \mathbf{Q} = \mathbf{1}$. The diagonalization of T is truncated at a dimension that is greatly reduced compared to the full dimension of possible excitations, generating an effective excitation spectrum. Using this approach the excitation spectrum converges to the exact excitation spectrum from the bottom or the top, i.e. higher energy starting vectors do converge on the higher energy.

Equation-of-Motion Coupled Cluster Singles and Doubles

Within the context of EOM-CCSD, it is convenient to define a normal ordered Hamiltonian (H_N) that can be defined as

$$\hat{H}_N = e^{-\hat{T}} \hat{H} e^{\hat{T}} - E, \quad (1.30)$$

this normal ordered Hamiltonian is then used to build the matrix elements of the EOM-CCSD Hamiltonian

$$\begin{aligned} \hat{H}_{\text{EOM-CCSD}} &= \begin{bmatrix} \hat{H}^{\text{SS}} & \hat{H}^{\text{SD}} \\ \hat{H}^{\text{DS}} & \hat{H}^{\text{DD}} \end{bmatrix} \\ \hat{H}^{\text{SS}} &= \langle \Phi_i^a | \hat{H}_N | \Phi_k^c \rangle \\ \hat{H}^{\text{SD}} &= \langle \Phi_i^a | \hat{H}_N | \Phi_{kl}^{cd} \rangle \\ \hat{H}^{\text{DS}} &= \langle \Phi_{ij}^{ab} | \hat{H}_N | \Phi_k^c \rangle \\ \hat{H}^{\text{DD}} &= \langle \Phi_{ij}^{ab} | \hat{H}_N | \Phi_{kl}^{cd} \rangle, \end{aligned} \quad (1.31)$$

where $|\Phi_i^a\rangle$ and $|\Phi_{ij}^{ab}\rangle$ are the singly and doubly excited determinants respectively. This EOM-CCSD Hamiltonian is then diagonalized using a modified Davidson algorithm to obtain the excitation energy eigenvalue spectrum. Similar to CC-LR, EOM-CCSD suffers from a feasibility problem as the high energy core-excited states can only be obtained after first obtaining a series of lower energy solutions. The need to solve for the full dimension of excitations in order to reach the core-excited states of interest causes standard diagonalization techniques to become computationally expensive for all cases and impossible for reasonably large cases. This issue can be handled in a

similar way to the energy specific TDDFT method outlined above, to yield an energy specific EOM method (ES-EOM-CCSD)⁹⁶

1.5 Orthogonality Constrained Density Functional Theory

In this section, I review orthogonality constrained density functional theory (OCDFT) as a general approach to electronic excited states. OCDFT belongs to a unique class of density functional approaches whose main goal is to study excited states without using a time-dependent formulation, otherwise known as time independent density functional theory (TIDFT).^{131,132} This class of methods aim to provide an alternative to TDDFT which has been known to struggle in cases of charge transfer and Rydberg excited states due to inadequacies of approximate exchange correlation functionals.^{133,134}

OCDFT is a variational formulation of DFT for the calculation of excited states.²⁹ Unique to the theory is the explicit enforcement of constraints that guarantee orthogonality while allowing the MOs to fully relax. This draws inspiration from two TIDFT theories:

- The variational TIDFT approach of Levy and Nagy which provides the theoretical framework for a universal excited state density functional,¹³⁵ thereby justifying the variational optimization of an excited state in DFT.
- The constrained DFT (CDFT) approach of Van Voorhis and coworkers¹³⁶ which provides a framework for studying systems within Kohn–Sham DFT that have an arbitrary constraint on their density.

The remainder of this section will provide a brief introduction to the theory, following this introduction will be a discussion of the features of the theory that specifically make it an attractive choice for the calculation of core-excited states.

1.5.1 Original Formulation via Constrained Variational Minimization

Within a generalized Kohn–Sham (KS) framework, an n -electron system can be represented in terms of an auxiliary system of noninteracting electrons.^{137,138} The KS auxiliary system can be represented by a single Slater determinant

$$|\Phi\rangle = |\phi_1, \phi_2, \dots, \phi_n\rangle, \quad (1.32)$$

which is constructed from a set of KS orbitals $\{\phi_i\}$. The total ground state electronic energy of the KS system can be expressed as

$$E_{\text{KS}}[\{\phi_i\}] = -\frac{1}{2} \sum_i^{\text{occ}} \langle \phi_i | \nabla^2 | \phi_i \rangle + \int d\mathbf{r} \nu(\mathbf{r}) \rho(\mathbf{r}) + J[\rho] + E'_{\text{XC}}[\rho], \quad (1.33)$$

where the first term is the kinetic energy contribution, the second term is the contribution of the external potential, while $J[\rho]$ and $E'_{\text{XC}}[\rho]$ are the Coulomb and exchange-correlation contribution respectively. In general, an excited state can be computed by a constrained minimization of the expectation value of the Hamiltonian (\hat{H}) over all the wave functions orthogonal to the ground state Ψ :

$$E' = \min_{\rho'} \left\{ \min_{\substack{\Psi' \rightarrow \rho' \\ \Psi' \perp \Psi}} \langle \Psi' | \hat{H} | \Psi' \rangle \right\} \quad (1.34)$$

$$= \min_{\rho'} \left\{ \int d\mathbf{r} \nu(\mathbf{r}) \rho(\mathbf{r}) + F'[\rho'] \right\}, \quad (1.35)$$

where $F'[\rho']$ minimizes the sum of the kinetic energy (\hat{T}) and electron-electron repulsion. In the spirit of the variational TIDFT of Levy and Nagy, OCDFT introduces a fictitious KS system of noninteracting electrons with density ρ'_s which maps to the excited state density of the real interacting system ρ' and can be described by a unique Slater determinant Φ' . The excited state solution is obtained via a *constrained minimization* subject to an orthogonality constraint on the auxiliary wavefunctions

$$\langle \Phi' | \Phi \rangle = 0, \quad (1.36)$$

Direct imposition of this orthogonality constraint avoids the excited state wavefunction from collapsing to the ground state solution. The OCDFT energy functional for the excited state is then expressed as

$$E'_{\text{OCDFT}}[\{\phi'_i\}] = -\frac{1}{2} \sum_i^{\text{occ}} \langle \phi'_i | \nabla^2 | \phi'_i \rangle + \int d\mathbf{r} \nu(\mathbf{r}) \rho'(\mathbf{r}) + J[\rho'] + E_{\text{XC}}[\rho'], \quad (1.37)$$

Note that the exchange-correlation functional E_{XC} is approximated with the ground state functional evaluated at the excited state density. This is an adiabatic approximation for the excited state and is discussed in more detail in Ref. 29.

1.5.2 Attractive Features of OCDFT for Core Excitations

This section will focus on the features of orthogonality constrained density functional theory that make it an attractive candidate as a theory for core excitations. Three features stand out the most and will be covered in more detail in this section:

- Full inclusion of orbital relaxation effects through variational optimization of Kohn–Sham orbitals.
- Variational stability of the Δ SCF procedure due to imposition of orbital orthogonality.
- Computational cost and scaling comparable to time-dependent density functional theory.

Having established itself as a stable, cheap, and accurate theory for electronic excited states the extension of this theory for the treatment of core excited states seemed like a logical progression.

Orbital Relaxation Effects

Within OCDFT, a general excited state determinant Φ' for an n -electron system is optimized with respect to a set of excited state orbitals $\{\phi'_n\}$

$$|\Phi'\rangle = |\phi'_1, \phi'_2, \dots, \phi'_n\rangle, \quad (1.38)$$

Since the excited state is variationally optimized, relaxation effects are explicitly considered for each excited state determinant. Within a quasiparticle framework, two of the orbitals that compose this Slater determinant will define the excitation process as the hole (ϕ'_h) and particle (ϕ'_p) respectively. The excited state orbitals are optimized subject to a set of minimal orbital orthogonality conditions: 1) The occupied orbitals are restricted to span the space complementary to the hole orbital

$$\langle \phi'_i | \phi'_h \rangle = 0, i = 1, 2, \dots, n \quad (1.39)$$

and 2) The hole orbital must be constrained to span the occupied space of the ground state determinant Φ . Optimizing the the excited state relative to these two constraints is sufficient to produce an orthogonal excited state solution. This formalism provides the flexibility to control the degree of optimization of the wavefunction, i.e. to control the amount of relaxation that it explicitly included in the calculation. This analysis was very useful for determining the importance of relaxation effects on the performance of OCDFT in valence excitations,²⁹ and a similar analysis is applied here to examine the importance for core excitations. We make use of 3 different levels of optimization 1) Obtaining the fully optimal hole/particle pair and allowing full relaxation of all alpha and beta orbitals 2) Obtaining the optimal hole/particle pair while only allowing relaxation of the alpha orbitals 3) and Obtaining the optimal hole orbital only while allowing all other alpha/beta orbitals to relax Figure 1.3 shows the results of applying these different levels of relaxation to the first core excited state in H₂O. Optimizing the hole and particle pair with full relaxation yields a total energy of -56.6582 Hartree. Only relaxing the α orbitals raises this energy by about 0.2 Hartree, showing that the rearrangement that occurs in the electronic structure due to the creation of a core hole is significant. This highlights the importance of valence rearrangement and inclusion of hole relaxation effects. The ability of OCDFT to explicitly account for this effect makes it an excellent method for computing core-excitation energies accurately.

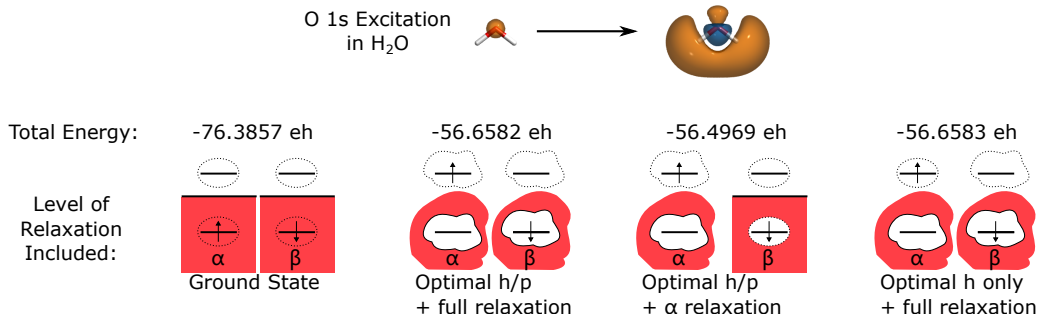


Figure 1.3: First O 1s core excitation in H₂O calculated using orthogonality constrained density functional theory at the B3LYP/3-21G level of theory. Numbers reported are the total energy of the ground state and respective core excited state calculated with varying degrees of orbital relaxation included.

Stability of the Self-Consistent Algorithm

The Ritz variational principle defines the energy expectation value \tilde{E} of any trial wavefunction $\tilde{\phi}$ as an upper-bound with respect to the ground state energy E_0 of the system

$$\tilde{E} = \frac{\langle \tilde{\phi} | \hat{H} | \tilde{\phi} \rangle}{\langle \tilde{\phi} | \tilde{\phi} \rangle} \geq E_0. \quad (1.40)$$

However, the energy of a variationally computed excited state E' is no longer a valid upper bound of the ground state energy unless the trial wavefunction ($\tilde{\phi}$) is constrained to be orthogonal to all lower lying states of the same symmetry. In this regard, a standard Δ SCF calculation of an excited state can be viewed as an unconstrained optimization, which in many cases guides the final optimized excited state solution to the ground state wave function. This is a well known optimization problem called “variational collapse.” Considering this issue, it is desirable to have a method that allows for the general variational calculation of an excited state without sacrificing the orthogonality of the computed excited state wavefunction to lower lying states of the same symmetry. Imposing orthogonality constraints into the variational procedure ensures the strict upperboundedness of the computed excited state energies, which ensures that the optimization will not suffer from variational collapse. This is the approach taken by orthogonality constrained density functional theory as

well as the “takes orthogonality constraints into account” (TOCIA) method.¹³⁹ Other methods such as Morokuma’s electron-hole potential (EHP) method¹⁴⁰ and the constrained variational DFT method (CV-DFT) of Ziegler et al.^{141,142} aim to avoid this optimization problem without imposing exact orthogonality and instead impose different variational constraints. Despite these differences, all of these methods revolve around the variational minimization a unique KS energy functional for the excited state determinant Φ' that is different from the ground state reference Φ . They do this explicitly by parametrizing the wavefunction using a generic determinant, this can be expressed as a unitary transformation of the ground state determinant

$$|\Phi'\rangle = |\phi'_1, \phi'_2, \dots, \phi'_n\rangle = e^{\hat{A}}|\Phi\rangle, \quad (1.41)$$

where \hat{A} is a one-particle anti-Hermitian operator. To impose an orthogonality condition on the excited state wavefunction ($\langle\Phi'|\Phi\rangle = 0$) OCDFT imposes the minimal orbital orthogonality constraints necessary to obtain an orthogonal solution. For example, to provide a description for an excited state, the excited state hole orbital (ϕ_h) is simply constrained to be orthogonal to the ground state occupied orbitals (ϕ_i)

$$\langle\phi'_h|\phi_i\rangle = 0, i = 1, 2, \dots, n \quad (1.42)$$

This constraint guarantees that the excited state determinant is orthogonal to the ground state, thus ensuring that the resulting optimization procedure does not suffer from variational collapse.

Favorable Computational Cost and Scaling

Cheap computational cost and favorable scaling with respect to system size are extremely important features of OCDFT that allows for a wide range of potential applications. Formally the cost of OCDFT is identical to that of a ground-state DFT computation which scales as N^3 for pure functionals and N^4 for hybrid functionals, where N is the number of electrons in the system. This scaling with respect to system

size is comparable to that of CIS and TDDFT which also scale as N^4 , in which the bottleneck in those methods is the contraction of a transition density matrix with the two-electron integrals. The scaling of OCDFE is also better than that of ADC(2) and excited state CC2 which scale as N^5 . Unlike CIS and TDDFT, OCDFE maintains this favorable cost while still obtaining accurate excitation energies comparable to more expensive methods, as will be shown in Chapters 2 and 3 of this dissertation.

1.6 Prospectus

In Chapter 2, we report an extension of orthogonality constrained DFT to compute core excited states and generalize the original formalism in order to calculate multiple excited state solutions. This chapter focuses on benchmarking the method with respect to experimental core excitation energies and simulating full experimental spectra. Chapter 3 reports the combination of OCDFE with a novel implementation of the spin-free exact-two-component (X2C) one-electron treatment of scalar relativistic effects in order to study the importance of relativistic effects on core-excited states and treat K-edges of transition metal complexes. Chapter 4 describes a new technique for the automated classification of core excited states that utilizes a unique representation of the unoccupied orbitals known as the localized intrinsic valence virtual orbitals (LIVVOs). This new method of classification is integrated with the OCDFE hole/particle orbital formalism here. Chapter 5 deals with the unique case of targeting the K-edge chemisorbed organic molecules. We introduce a subspace projection algorithm into OCDFE in order to specifically target the core excitations of the organic adsorbate, which is not feasible in the bottom-up algorithm described in Chapter 2.

Bibliography

- [1] Marguet, S.; Markovitsi, D.; Millié, P.; Sigal, H.; Kumar, S. *J. Phys. Chem. B* **1998**, *102*, 4697–4710.

- [2] González, L.; Escudero, D.; Serrano-Andrés, L. *ChemPhysChem* **2012**, *13*, 28–51.
- [3] Zhou, Q. L.; Heflin, J. R.; Wong, K. Y.; Zamani-Khamiri, O.; Garito, A. F. *Phys. Rev. A* **1991**, *43*, 1673–1676.
- [4] Hirata, S.; Head-Gordon, M.; Bartlett, R. J. *J. Chem. Phys.* **1999**, *111*, 10774–10786.
- [5] Harris, D. C.; Bertolucci, M. D. *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy*; Courier Corporation, 1978.
- [6] Serrano-Andrés, L.; Fülcher, M. P. Theoretical Study of the Electronic Spectroscopy of Peptides. 1. The Peptidic Bond: Primary, Secondary, and Tertiary Amides - Journal of the American Chemical Society (ACS Publications). 1998; <http://pubs.acs.org.proxy.library.emory.edu/doi/full/10.1021/ja961996+>.
- [7] Roos, B. O.; Andersson, K.; Fülcher, M. P.; Malmqvist, P.-å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons, Inc., 1996; pp 219–331, DOI: 10.1002/9780470141526.ch5.
- [8] Autschbach, J. *ChemPhysChem* **2009**, *10*, 1757–1760.
- [9] Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol. Phys.* **2006**, *104*, 1039–1051.
- [10] Prlj, A.; Sandoval-Salinas, M. E.; Casanova, D.; Jacquemin, D.; Corminboeuf, C. *J. Chem. Theory Comput.* **2016**, *12*, 2652–2660.
- [11] Nakata, A.; Imamura, Y.; Otsuka, T.; Nakai, H. *J. Chem. Phys.* **2006**, *124*, 094105.
- [12] Penner-Hahn, J. E. *Coord. Chem. Rev.* **1999**, *190192*, 1101–1123.
- [13] Stöhr, J. *NEXAFS Spectroscopy*; Springer: Berlin, 1992; Vol. 25.
- [14] Snigirev, A.; Kohn, V.; Snigireva, I.; Lengeler, B. *Nature; London* **1996**, *384*, 49–51.
- [15] Holch, F.; Hübner, D.; Fink, R.; Schöll, A.; Umbach, E. *J. Electron. Spectrosc. Relat. Phenom.* **2011**, *184*, 452–456.
- [16] Lee, P. A.; Citrin, P. H.; Eisenberger, P.; Kincaid, B. M. *Rev. Mod. Phys.* **1981**, *53*, 769–806.
- [17] Rehr, J. J.; Albers, R. C. *Rev. Mod. Phys.* **2000**, *72*, 621–654.
- [18] Snyder, L. C. CoreElectron Binding Energies and Slater Atomic Shielding Constants: The Journal of Chemical Physics: Vol 55, No 1. 1971; <http://aip.scitation.org.proxy.library.emory.edu/doi/abs/10.1063/1.1675565>.
- [19] Schirmer, J.; Braunstein, M.; Lee, M.-T.; McKoy, V. In *VUV and Soft X-Ray Photoionization*; Becker, U., Shirley, D. A., Eds.; Physics of Atoms and Molecules; Springer US, 1996; pp 105–133, DOI: 10.1007/978-1-4613-0315-2_4.
- [20] Decleva, P.; Fronzoni, G.; Lisini, A.; Stener, M. *Chem. Phys.* **1994**, *186*, 1–16.
- [21] Ankudinov, A. L.; Rehr, J. J. *Phys. Rev. B* **1997**, *56*, R1712–R1716.
- [22] Kutzler, F. W.; Natoli, C. R.; Misemer, D. K.; Doniach, S.; Hodgson, K. O. *J. Chem. Phys.* **1980**, *73*, 3274–3288.
- [23] Bagus, P. S.; Weiss, K.; Schertel, A.; Wöll, C.; Braun, W.; Hellwig, C.; Jung, C. *Chem. Phys. Lett.* **1996**, *248*, 129–135.

- [24] Glaser, T.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *Acc. Chem. Res.* **2000**, *33*, 859–868.
- [25] Urquhart, S. G.; Ade, H. *J. Phys. Chem. B* **2002**, *106*, 8531–8538.
- [26] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2011**, *134*, 244304.
- [27] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2013**, *138*, 214302.
- [28] Behyan, S.; Hu, Y.; Urquhart, S. G. *Chem. Phys. Lett.* **2014**, *592*, 69–74.
- [29] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- [30] Holick, M. F. *J. Cell. Biochem.* **2003**, *88*, 296–307.
- [31] Krause, G. H.; Weis, E. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **1991**, *42*, 313–349.
- [32] de Vries, H. L. *Physica* **1943**, *10*, 553–564.
- [33] Franck, J.; Dymond, E. G. *Trans. Faraday Soc.* **1926**, *21*, 536–542.
- [34] Condon, E. *Phys. Rev.* **1926**, *28*, 1182–1201.
- [35] Condon, E. U. *Phys. Rev.* **1928**, *32*, 858–872.
- [36] Hazra, A.; Chang, H. H.; Nooijen, M. *J. Chem. Phys.* **2004**, *121*, 2125–2136.
- [37] Maxwell, K.; Johnson, G. N. *J Exp Bot* **2000**, *51*, 659–668.
- [38] Valeur, B.; Berberan-Santos, M. N. *Molecular Fluorescence: Principles and Applications*; John Wiley & Sons, 2012; Google-Books-ID: NdySsFt8TN8C.
- [39] Tanaka, I.; Tabata, Y.; Tokito, S. *Phys. Rev. B* **2005**, *71*, 205207.
- [40] Romanovskii, Y. V.; Gerhard, A.; Schweitzer, B.; Scherf, U.; Personov, R. I.; Bässler, H. *Phys. Rev. Lett.* **2000**, *84*, 1027–1030.
- [41] avar, E.; Blüm, M.-C.; Pivetta, M.; Patthey, m. . n. y. . . p. . . f. . A., François and Chergui, Majed and Schneider, Wolf-Dieter *Phys. Rev. Lett.* *95*.
- [42] Callomon, J. H.; Parkin, J. E.; Lopez-Delgado, R. *Chem. Phys. Lett.* **1972**, *13*, 125–131.
- [43] De Mello Donegá, C.; Meijerink, A.; Blasse, G. *J. Phys. Chem. Solids* **1995**, *56*, 673–685.
- [44] Lamola, A. A.; Hammond, G. S. *J. Chem. Phys.* **1965**, *43*, 2129–2135.
- [45] Hauser, A. *Coord. Chem. Rev.* **1991**, *111*, 275–290.
- [46] Zakrzewski, J.; Rzazewski, K.; Lewenstein, M. *J. Phys. B: At. Mol. Phys.* **1984**, *17*, 729.
- [47] Meskers, S. C. J.; Janssen, R. A. J.; Haverkort, J. E. M.; Wolter, J. H. *Chem. Phys.* **2000**, *260*, 415–439.
- [48] Van't Hof, C. A.; Schmidt, J. *Chem. Phys. Lett.* **1976**, *42*, 73–77.
- [49] Hors, P. J.; Egmond, M. R.; Edzes, H. T.; Kaptein, R. *J Magn Reson.* **1982**, *49*, 122–150.
- [50] Politzer, P.; Parr, R. G. *J. Chem. Phys.* **1976**, *64*, 4634–4637.

- [51] Yoon, W.-S.; Kim, K.-B.; Kim, M.-G.; Lee, M.-K.; Shin, H.-J.; Lee, J.-M.; Lee, J.-S.; Yo, C.-H. *J. Phys. Chem. B* **2002**, *106*, 2526–2532.
- [52] Parent, P.; Laffon, C.; Mangeney, C.; Bournel, F.; Tronc, M. *J. Chem. Phys.* **2002**, *117*, 10842–10851.
- [53] Lambrecht, W. R. L.; Rashkeev, S. N.; Segall, B.; Lawniczak-Jablonska, K.; Suski, T.; Gullikson, E. M.; Underwood, J. H.; Perera, R. C. C.; Rife, J. C.; Grzegory, I.; Porowski, S.; Wickenden, D. K. *Phys. Rev. B* **1997**, *55*, 2612–2622.
- [54] Leinweber, P.; Kruse, J.; Walley, F. L.; Gillespie, A.; Eckhardt, K.-U.; Blyth, R. I. R.; Regier, T. *J. Synchrotron Radiat* **2007**, *14*, 500–511.
- [55] Cotte, M.; Susini, J.; Dik, J.; Janssens, K. *Acc. Chem. Res.* **2010**, *43*, 705–714.
- [56] Yamamoto, T. *X-Ray Spectrom.* **2008**, *37*, 572–584.
- [57] Vedrinskii, R. V.; Kraizman, V. L.; Novakovich, A. A.; Demekhin, P. V.; Urazhdin, S. V. *J. Phys.: Condens. Matter* **1998**, *10*, 9561.
- [58] Groot, F. d.; Vankó, G.; Glatzel, P. *J. Phys.: Condens. Matter* **2009**, *21*, 104207.
- [59] Roemelt, M.; Beckwith, M. A.; Duboc, C.; Collomb, M.-N.; Neese, F.; DeBeer, S. *Inorg. Chem.* **2012**, *51*, 680–687.
- [60] Lengeler, B.; Eisenberger, P. *Phys. Rev. B* **1980**, *21*, 4507–4520.
- [61] Teo, B. K. *EXAFS: Basic Principles and Data Analysis*; Springer Science & Business Media, 2012; Vol. 9.
- [62] Gurman, S. J.; Binsted, N.; Ross, I. *J. Phys. C: Solid State Phys.* **1984**, *17*, 143.
- [63] Bunker, G. *Nucl. Instr. Meth. Phys. Res. B* **1983**, *207*, 437–444.
- [64] Garino, C.; Borfecchia, E.; Gobetto, R.; van Bokhoven, J. A.; Lamberti, C. *Coord. Chem. Rev.* **2014**, *277278*, 130–186.
- [65] Tourillon, G.; Mahatsekake, C.; Andrieu, C.; Williams, G. P.; Garrett, R. F.; Braun, W. *Surf. Sci.* **1988**, *201*, 171–184.
- [66] Zubavichus, Y.; Zharnikov, M.; Schaporenko, A.; Grunze, M. *J. Electron. Spectrosc. Relat. Phenom.* **2004**, *134*, 25–33.
- [67] Zubavichus, Y.; Shaporenko, A.; Grunze, M.; Zharnikov, M. *Nucl. Instr. Meth. Phys. Res. A* **2009**, *603*, 111–114.
- [68] Samuel, N. T.; Fischer, D. A.; Castner, D. G. *Surf. Interface. Anal.* **2017**,
- [69] Contini, G.; Carravetta, V.; Di Castro, V.; Stranges, S.; Richter, R.; Alagia, M. *J. Phys. Chem. A* **2001**, *105*, 7308–7314.
- [70] Domke, M.; Xue, C.; Puschmann, A.; Mandel, T.; Hudson, E.; Shirley, D. A.; Kaindl, G. *Chem. Phys. Lett.* **1990**, *173*, 122–128.
- [71] Hudson, E.; Shirley, D. A.; Domke, M.; Remmers, G.; Kaindl, G. *Phys. Rev. A* **1994**, *49*, 161–175.

- [72] Roy, S. S.; McCann, R.; Papakonstantinou, P.; Maguire, P.; McLaughlin, J. A. *Thin Solid Films* **2005**, *482*, 145–150.
- [73] Hahner, G. *Chem. Soc. Rev.* **2006**, *35*, 1244–1255.
- [74] Baldo, M. A.; O'Brien, D. F.; Thompson, M. E.; Forrest, S. R. *Phys. Rev. B* **1999**, *60*, 14422–14428.
- [75] Guo, Q. X.; Senda, H.; Saito, K.; Tanaka, T.; Nishio, M.; Ding, J.; Fan, T. X.; Zhang, D.; Wang, X. Q.; Liu, S. T.; Shen, B.; Ohtani, R. *Appl. Phys. Lett.* **2011**, *98*, 181901.
- [76] Hoffmann, H.; Zaera, F.; Mark Ormerod, R.; Lambert, R. M.; Wang, L. P.; Tysoe, W. T. *Surf. Sci.* **1990**, *232*, 259–265.
- [77] Stöhr, J.; Gland, J. L.; Eberhardt, W.; Outka, D.; Madix, R. J.; Sette, F.; Koestner, R. J.; Doebler, U. *Phys. Rev. Lett.* **1983**, *51*, 2414–2417.
- [78] Temprano, I.; Thomas, G.; Haq, S.; Dyer, M. S.; Latter, E. G.; Darling, G. R.; Uvdal, P.; Raval, R. *J. Chem. Phys.* **2015**, *142*, 101916.
- [79] Rastgoo-Lahrood, A.; Björk, J.; Lischka, M.; Eichhorn, J.; Kloft, S.; Fritton, M.; Strunskus, T.; Samanta, D.; Schmittel, M.; Heckl, W. M.; Lackinger, M. *Angew. Chem. Int. Ed.* **2016**, *55*, 7650–7654.
- [80] Lee, H.-K.; Park, J.; Kim, I.; Kim, H.-D.; Park, B.-G.; Shin, H.-J.; Lee, I.-J.; Singh, A. P.; Thakur, A.; Kim, J.-Y. *J. Phys. Chem. C* **2012**, *116*, 722–725.
- [81] Chen, L. X. *Annu. Rev. Phys. Chem.* **2005**, *56*, 221–254.
- [82] Attar, A. R.; Bhattacharjee, A.; Pemmaraju, C. D.; Schnorr, K.; Closser, K. D.; Prendergast, D.; Leone, S. R. *Science* **2017**, *356*, 54–59.
- [83] Day, A. C. *J. Am. Chem. Soc.* **1975**, *97*, 2431–2438.
- [84] Baeck, K. K.; Martinez, T. J. *Chem. Phys. Lett.* **2003**, *375*, 299–308.
- [85] Gwaltney, S. R.; Bartlett, R. J. *J. Chem. Phys.* **1998**, *108*, 6790–6798.
- [86] Shih, S.-k.; Peyerimhoff, S. D.; Buenker, R. J. *Chem. Phys.* **1976**, *17*, 391–402.
- [87] Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.
- [88] Fantacci, S.; De Angelis, F.; Selloni, A. *J. Am. Chem. Soc.* **2003**, *125*, 4381–4387.
- [89] De Angelis, F.; Fantacci, S.; Sgamellotti, A.; Cariati, E.; Ugo, R.; Ford, P. C. *Inorg. Chem.* **2006**, *45*, 10576–10584.
- [90] Gwaltney, S. R.; Nooijen, M.; Bartlett, R. J. *Chem. Phys. Lett.* **1996**, *248*, 189–198.
- [91] Roos, B. *Chem. Phys. Lett.* **1972**, *15*, 153–159.
- [92] Nakatsuji, H.; Davidson, E. R. *J. Chem. Phys.* **2001**, *115*, 2000–2006.
- [93] Mazur, G.; Wodarczyk, R. *J. Comput. Chem.* **2009**, *30*, 811–817.
- [94] Imamura, Y.; Nakai, H. *Chem. Phys. Lett.* **2006**, *419*, 297–303.

- [95] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024.
- [96] Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.
- [97] Lestrange, P. J.; Nguyen, P. D.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.
- [98] Ehlert, C.; Klamroth, T. *J. Comput. Chem.* **2017**, *38*, 116–126.
- [99] Wenzel, J.; Wormit, M.; Dreuw, A. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.
- [100] Triguero, L.; Plashkevych, O.; Pettersson, L. G. M.; Ågren, H. *J. Electron. Spectrosc. Relat. Phenom.* **1999**, *104*, 195–207.
- [101] In Koopmans theorem the i^{th} Ionization potential is approximated by the corresponding energy eigenvalue of the canonical SCF equation $-\epsilon_i$.
- [102] Plakhutin, B. N.; Gorelik, E. V.; Breslavskaya, N. N. *J. Chem. Phys.* **2006**, *125*, 204110.
- [103] Reiher, M. *Theor Chem Acc* **2006**, *116*, 241–252.
- [104] Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. M. *Chem. Phys. Lett.* **1994**, *222*, 75–81.
- [105] Hunt, s. J.; Goddard III, W. A. *Chem. Phys. Lett.* **1969**, *3*, 414–418.
- [106] Zhang, Y.; Hua, W.; Bennett, K.; Mukamel, S. In *Density-Functional Methods for Excited States*; Ferré, N., Filatov, M., Huix-Rotllant, M., Eds.; Topics in Current Chemistry 368; Springer International Publishing, 2014; pp 273–345, DOI: 10.1007/128_2014_618.
- [107] Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. M. *Theor Chem Acta* **1997**, *97*, 14–40.
- [108] Besley, N. A.; Noble, A. *J. Phys. Chem. C* **2007**, *111*, 3333–3340.
- [109] Fransson, T.; Coriani, S.; Christiansen, O.; Norman, P. *J. Chem. Phys.* **2013**, *138*, 124311.
- [110] Besley, N. A. *Chem. Phys. Lett.* **2012**, *542*, 42–46.
- [111] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A* **2012**, *85*, 022507.
- [112] Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. *J. Chem. Theory Comput.* **2012**, *8*, 1616–1628.
- [113] Myhre, R. H.; Coriani, S.; Koch, H. *J. Chem. Theory Comput.* **2016**, *12*, 2633–2643.
- [114] Wenzel, J.; Holzer, A.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2015**, *142*, 214104.
- [115] Wenzel, J.; Dreuw, A. *J. Chem. Theory Comput.* **2016**, *12*, 1314–1330.
- [116] Mochizuki, Y.; Koide, H.; Imamura, T.; Takemiya, H. *J. Synchrotron Radiat.* **2001**, *8*, 1003–1005.
- [117] Maganas, D.; Kristiansen, P.; Duda, L.-C.; Knop-Gericke, A.; DeBeer, S.; Schlögl, R.; Neese, F. *J. Phys. Chem. C* **2014**, *118*, 20163–20175.
- [118] Maganas, D.; Roemelt, M.; Weyhermuller, T.; Blume, R.; Havecker, M.; Knop-Gericke, A.; DeBeer, S.; Schlogl, R.; Neese, F. *Phys. Chem. Chem. Phys.* **2014**, *16*, 264–276.

- [119] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.
- [120] Asmurf, F.; Besley, N. *Chem. Phys. Lett.* **2008**, *463*, 267–271.
- [121] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.
- [122] Olovsson, W.; Tanaka, I.; Puschnig, P.; Ambrosch-Draxl, C. *J. Phys.: Condens. Matter* **2009**, *21*, 104205.
- [123] Vinson, J. *Phys. Rev. B* **2012**, *85*.
- [124] Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 204101.
- [125] Ebert, H.; Stöhr, J.; Parkin, S. S. P.; Samant, M.; Nilsson, A. *Phys. Rev. B* **1996**, *53*, 16067–16073.
- [126] Maganas, D.; Roemelt, M.; Hävecker, M.; Trunschke, A.; Knop-Gericke, A.; Schlögl, R.; Neese, F. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7260–7276.
- [127] Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.
- [128] Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. *J. Chem. Phys.* **2009**, *130*, 124308.
- [129] Fetter, A.; Walecka, J. D. *Quantum Theory of Many-Particle Systems*; McGraw-Hill: New York, 1971.
- [130] Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- [131] Samal, P.; Harbola, M. K. *J. Phys. B: At. Mol. Opt. Phys.* **2006**, *39*, 4065.
- [132] Ayers, P. W.; Levy, M. *Phys. Rev. A* **2009**, *80*, 012508.
- [133] Zhang, Y.; Yang, W. *J. Chem. Phys.* **1998**, *109*, 2604–2608.
- [134] Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Science* **2008**, *321*, 792–794.
- [135] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A* **2012**, *85*, 042518.
- [136] Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. *Chem. Rev.* **2012**, *112*, 321–370.
- [137] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- [138] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [139] Tsaune, A. Y.; Glushkov, V. N.; Aprasyukhin, A. I. *THEOCHEM* **1994**, *312*, 289–295.
- [140] Morokuma, K.; Iwata, S. *Chem. Phys. Lett.* **1972**, *16*, 192–197.
- [141] Cullen, J.; Krykunov, M.; Ziegler, T. *Chemical Physics* **2011**, *391*, 11–18.
- [142] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Phys.* **2012**, *136*, 124107.

Chapter 2

Simulation of X-Ray Absorption Spectra with Orthogonality Constrained Density Functional Theory

“The history of science is rich in the example of the fruitfulness of bringing two sets of techniques, two sets of ideas, developed in separate contexts for the pursuit of new truth, into touch with one another.”

J. Robert Oppenheimer

Chapter Abstract

Orthogonality constrained density functional theory (OCDFT) [F. A. Evangelista, P. Shushkov and J. C. Tully, *J. Phys. Chem. A*, 2013, **117**, 7378] is a variational time-independent approach for the computation of electronic excited states. In this work we extend OCDFT to compute core-excited states and generalize the original formalism to determine multiple excited states. Benchmark computations on a set of 13 small molecules and 40 excited states show that *unshifted* OCDFT/B3LYP excitation energies have a mean absolute error of 1.0 eV. Contrary to time-dependent DFT, OCDFT excitation energies for first- and second-row elements are computed with near-uniform accuracy. OCDFT core excitation energies are insensitive to the choice of the functional and the amount of Hartree–Fock exchange. We show that

OCDFT is a powerful tool for the assignment of X-ray absorption spectra of large molecules by simulating the gas-phase near-edge spectrum of adenine and thymine.

2.1 Introduction

The advent of synchrotron light sources created a strong resurgence of spectroscopy in the X-ray region.¹ Near-edge X-ray absorption spectroscopy (NEXAS) is a useful experimental technique to probe the local electronic and geometrical structure in a variety of molecular environments. The most dominant feature of NEXAS spectra, the *near edge* (see Fig. 2.1), is composed of the excitations of core-electrons to unoccupied valence orbitals. Core excitations are atom-specific and sensitive to the local chemical environment, thus, NEXAS spectra can provide information about the chemical composition and the electronic structure of molecules. NEXAS has been successfully applied to large biological systems,² small molecules in the gas phase,³ organic thin-films,⁴ and semiconducting materials.⁵ This wide range of applications is possible because synchrotron light sources can span an energy range that goes from a few electron volt (eV)⁶ to hundreds of MeV.⁷

As NEXAS experiments are becoming more feasible, there is a growing need to develop accurate theoretical approaches to aid the interpretation of experimen-

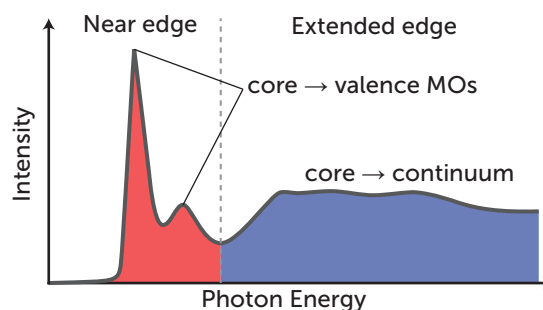


Figure 2.1: Example of a X-ray photoabsorption spectrum (XAS). The near edge, located in the low energy region, consists of excitations of core electrons to valence orbitals. These transitions are sensitive to the chemical environment surrounding the excited atom. The high energy region of the spectrum results from excitations of core electrons to the continuum.

tal spectra. Calculations of NEXAS spectra are challenging, and require computational methods that explicitly account for the excitations of core-level electrons, orbital relaxation effects, and electron correlation.⁸ Several theoretical approaches have been adapted to compute core-valence excitations, including: scaled-opposite-spin configuration interaction singles with perturbative doubles [SOS-CIS(D)],⁹ a restricted open-shell DFT/CIS method,^{10,11} second-order algebraic digrammatic construction [ADC(2)],^{12,13} multiple scattering X_α methods,¹⁴ a maximum overlap Δ SCF approach,¹⁵ restricted active space SCF method (RASSCF),¹⁶ transition potential theory,¹⁷ coupled-cluster response theory,⁸ time-dependent density functional theory (TDDFT),¹⁸ and restricted excitation window TDDFT (REW-TDDFT).¹⁹ Among these methods, TDDFT is perhaps the most attractive option because of its reduced computational cost and ability to calculate multiple excited states.

TDDFT is a rigorous extension of the DFT ground-state formalism,²⁰ and it is regarded as the method of choice to treat electronic excited states within a density functional framework. When applied in conjunction with frequency-independent exchange-correlation potentials, TDDFT yields accurate excitation energies for low-lying excited states. For example, the TDDFT benchmark study of Silva-Junior and co-workers²¹ on 28 organic molecules, showed that singlet and triplet excitation energies can be calculated with a mean average error (MAE) of 0.27 eV and 0.44 eV, respectively. However, Besley et al.¹⁵ showed that TDDFT core excitations computed with conventional exchange-correlation functionals grossly underestimate experimental results, yielding a MAE of 20.2 eV. It is customary to remedy this deficiency of TDDFT by shifting the position of computed spectra by an amount that minimizes the difference between the computational and experimental peak features. For example, a study done by DeBeer, Petrenko, and Neese²² showed that it is necessary shift the TDDFT Fe near-edge spectrum of different iron complexes by 171.3 eV in order to achieve quantitative agreement with experiment. This shift is dependent on the

basis set and functional used and must be recalibrated for every basis set/functional combination applied to a given system.

Work by Peach et al.²³ provided evidence for a direct correlation between the accuracy of TDDFT and the degree of spatial overlap (Λ) between the occupied and virtual orbitals. The inaccuracy of TDDFT for excitations with low values of Λ has been attributed to the incorrect asymptotic behavior of the exchange-correlation potential and self-interaction error.²³ Low orbital overlap is a characteristic feature of both charge-transfer excitations and core excitations, and as expected, the same correlation between accuracy and orbital overlap is observed for core excitations.²⁴ As a consequence, TDDFT core-valence excitation energies can be improved by introducing self-interaction corrections (SIC)²⁵ or range separated hybrid functionals in which the amount of long and short range Hartree–Fock exchange is reparametrized.^{24,26} However, the optimal value of the range separation parameter and the amount of Hartree–Fock exchange are strongly system dependent and must be tuned.^{27–29}

A general method that can systematically produce accurate core-excitation energies with traditional (that is, non range corrected) hybrid density functionals is highly desirable. The maximum overlap method (MOM)¹⁵ combined with a Δ SCF treatment of core-valence excitations is able to obtain highly accurate excitation energies using conventional functionals. However, this procedure is not guaranteed to avoid the problem of variational collapse—albeit MOM ameliorates the difficulties encountered by a straightforward Δ SCF procedure—and has not been generalized to multiple excited states of the same symmetry.

The goal of this work is to find cost-effective alternative theories to TDDFT that can be used to simulate NEXAS spectra. Orthogonality constrained density functional theory (OCDFT)³⁰ was rigorously derived from a variational time-independent formulation of excited state DFT. It builds upon previous successful efforts to formulate variational excited state DFT, such as: the Δ SCF procedure,^{31,32} constrained

DFT,³³ stationary state DFT,³⁴ constricted variational density functional theory (CV-DFT),^{35–38} perturbative constrained excited state DFT,^{39–41} ensemble DFT,^{42–45} and variational time independent DFT (TI-DFT).^{46,47} Formally, OCDFT may be viewed as bridging constrained and constricted variational DFT. Its main advantages are: 1) a favorable accuracy/cost ratio, similar to that of ground state DFT, 2) a numerically robust optimization procedure that avoids variational collapse, and 3) the ability to compute spin adapted excitation energies. Benchmark computations³⁰ show that valence excitation energies computed with OCDFT have error metrics comparable to that of TDDFT. In addition, OCDFT has the ability to accurately compute charge-transfer excitation energies regardless of the amount of Hartree–Fock exchange present in the exchange–correlation functional.

This work introduces two new developments of OCDFT that are necessary for the simulation of near-edge X-ray absorption spectra. First, we formulate an OCDFT algorithm that can be used to compute core-valence excitation energies. This new method is assessed over a test set that includes 13 molecules with 40 unique core-electron excitations. Second, we discuss one approach to extend OCDFT to multiple excited states of the same symmetry. We demonstrate the potential of this new method with computations of the gas-phase near-edge spectrum of adenine and thymine.

2.2 Theory

In this section we provide a brief summary of orthogonality constrained density functional theory along with the necessary extension to multiple excited states (for the full details of the OCDFT derivation we refer the reader to Ref. 30). OCDFT builds upon the time-independent variational DFT approach developed by Ayers, Levy, and Nagy.⁴⁸ Within this framework, each electronic state ($\Psi^{(n)}, n = 0, 1, \dots$) of a N -electron system has a corresponding density functional ($E^{(n)}[\rho]$) which is a gener-

alization of the ground-state functional of Levy. $E^{(n)}[\rho]$ minimizes the energy expectation value $\langle \Psi | \hat{H} | \Psi \rangle$ imposing two constraints on to the trial wave function Ψ : 1) Ψ must be compatible with the density ρ , and 2) Ψ must be orthogonal to the first $n - 1$ *exact* electronic states, $\{\Psi^{(k)}, k = 1, \dots, n - 1\}$:

$$E^{(n)}[\rho] = \min_{\substack{\Psi \rightarrow \rho \\ \Psi \perp \{\Psi^{(k)}\}}} \langle \Psi | \hat{H} | \Psi \rangle. \quad (2.1)$$

OCDFT provides a practical realization of this time-independent DFT approach. The first step in the OCDFT derivation consists in defining a generalized Kohn–Sham scheme that, for each electronic state $\Psi^{(n)}$, postulates an auxiliary system of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho_s^{(n)}$. The density of state $\Phi^{(n)}$ is assumed to be equal to the density of the exact state $\Psi^{(n)}$, which we indicate with $\rho^{(n)}$. In addition, we assume that the auxiliary wave functions are orthogonal, that is:

$$\langle \Phi^{(m)} | \Phi^{(n)} \rangle = \delta_{mn}. \quad (2.2)$$

This condition can be imposed without loss of generality. It avoids the excited state wave functions from collapsing down to the ground state solution, and effectively transfers some of the complexity of the excited state density functionals to the kinetic energy operator. Nevertheless, this variational Kohn–Sham scheme involves energy functionals $[E_{\text{KS}}^{(n)}]$ that contain exchange-correlation contributions $[E_{\text{xc}}^{(n)}]$ *specific* for each excited state. In OCDFT, we invoke an adiabatic approximation similar to the one used in TDDFT, and replace $E_{\text{xc}}^{(n)}$ with the ground state exchange-correlation functional, $E_{\text{xc}}^{(0)}$. The resulting functional for excited state n is given by:

$$E_{\text{OCDFT}}^{(n)}[\{\phi_i^{(n)}\}] = -\frac{1}{2} \sum_i^{\text{occ}} \langle \phi_i^{(n)} | \nabla^2 | \phi_i^{(n)} \rangle + \int d\mathbf{r} v(\mathbf{r}) \rho^{(n)}(\mathbf{r}) + J[\rho^{(n)}] + E_{\text{xc}}^{(0)}[\rho^{(n)}]. \quad (2.3)$$

Minimization of the $E^{(n)}[\{\phi_i^{(n)}\}]$ with respect to the occupied orbitals for state n can be performed with a modified self-consistent-field algorithm.³⁰ In the case of the

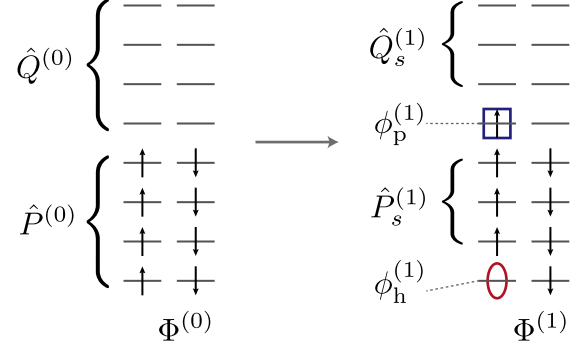


Figure 2.2: Illustration of the projection operators used in OCDFT. Notice, that a consequence of Eqs. (2.4)–(2.5) is that the hole $[\phi_h^{(1)}]$ and particle $[\phi_p^{(1)}]$ orbitals are contained in the spaces spanned by $\hat{P}^{(0)}$ and $\hat{Q}^{(0)}$.

first excited ($n = 1$), it is possible to show that the orthogonality condition [Eq. (2.2)] implies the existence of two special orbitals. As illustrated in Fig. 2.2, these are the *hole* $[\phi_h^{(1)}]$ and *particle* $[\phi_p^{(1)}]$ orbitals, which are respectively unoccupied and occupied in the excited state wave function $\Phi^{(1)}$. These orbitals must satisfy the conditions:

$$\hat{Q}^{(0)}\phi_h^{(1)} = 0, \quad (2.4)$$

$$\hat{P}^{(0)}\phi_p^{(1)} = 0, \quad (2.5)$$

where $\hat{P}^{(0)} = \sum_i |\phi_i^{(0)}\rangle\langle\phi_i^{(0)}|$ is a projector onto the occupied orbitals of $\Phi^{(0)}$, and $\hat{Q}^{(0)} = 1 - \hat{P}^{(0)}$. Eqs. (2.4) and (2.5) can be enforced via Lagrangian multipliers. Setting the variation of the Lagrangian with respect to the occupied ($\{\phi_i^{(1)}\}$), hole ($\phi_h^{(1)}$), and particle ($\phi_p^{(1)}$) orbitals to zero gives the following eigenvalue equations:

$$(1 - \hat{P}_{h/p}^{(1)})\hat{f}^{(1)}(1 - \hat{P}_{h/p}^{(1)})|\phi_i^{(1)}\rangle = \epsilon_i^{(1)}|\phi_i^{(1)}\rangle, \quad (2.6)$$

$$\hat{P}^{(0)}(1 - \hat{Q}_s^{(1)})\hat{f}^{(1)}(1 - \hat{Q}_s^{(1)})\hat{P}^{(0)}|\phi_h^{(1)}\rangle = \epsilon_h^{(1)}|\phi_h^{(1)}\rangle, \quad (2.7)$$

$$\hat{Q}^{(0)}(1 - \hat{P}_s^{(1)})\hat{f}^{(1)}(1 - \hat{P}_s^{(1)})\hat{Q}^{(0)}|\phi_p^{(1)}\rangle = \epsilon_p^{(1)}|\phi_p^{(1)}\rangle, \quad (2.8)$$

where $\hat{f}^{(1)}$ is the Kohn–Sham Hamiltonian of the excited state. Eq. (2.6) determines the occupied orbitals, while Eqs. (2.7) and (2.8) determine the hole and particle orbitals, respectively. The projection operators involved in the OCDFT equations

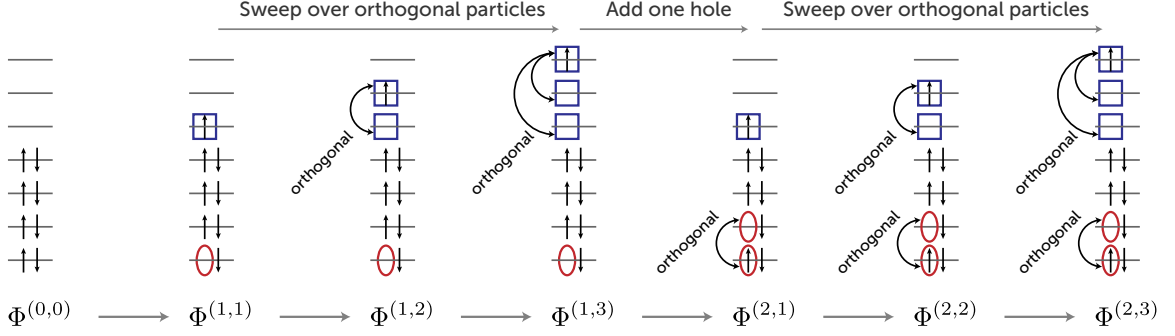


Figure 2.3: Constrained multiple hole/particle (CMHP) algorithm illustrated in the case of two hole and three particle orbitals.

are defined as (see Fig. 2.2):

$$\hat{P}_{\text{h/p}}^{(1)} = \hat{P}_{\text{h}}^{(1)} + \hat{P}_{\text{p}}^{(1)} = |\phi_{\text{h}}^{(1)}\rangle\langle\phi_{\text{h}}^{(1)}| + |\phi_{\text{p}}^{(1)}\rangle\langle\phi_{\text{p}}^{(1)}|, \quad (2.9)$$

$$\hat{P}_{\text{s}}^{(1)} = \hat{P}^{(1)} - \hat{P}_{\text{p}}^{(1)}, \quad (2.10)$$

$$\hat{Q}_{\text{s}}^{(1)} = \hat{Q}^{(1)} - \hat{P}_{\text{h}}^{(1)}, \quad (2.11)$$

where the subscript “s” stands for *spectator* orbitals. In OCDFT computations of valence excited states, the hole orbital is assumed to be the solution of Eq. (2.7) with the highest value of $\epsilon_{\text{h}}^{(1)}$. Similarly, the particle orbital corresponds to the lowest eigenvalue of Eq. (2.8).

Our OCDFT approach for core-excited states introduces two new aspects. First, in order to compute core-excited states with OCDFT, we select hole orbitals with the smallest values of $\epsilon_{\text{h}}^{(1)}$. However, this simple extension allows us only to compute one core-excited state for each irreducible representation. In principle, OCDFT can be generalized to compute an arbitrary number of excited states. For each additional excited state, it is necessary to minimize the OCDFT energy imposing orthogonality with respect to the lower energy solutions. While this appears to be a viable solution, it would undoubtedly lead to a more elaborate minimization procedure. In this work we propose simplified orthogonality conditions that are based on the orthogonality of the hole and/or particle orbitals. For example, if we choose the hole orbital for the second excited state $[\phi_{\text{h}}^{(2)}]$ to be orthogonal to the first hole $[\phi_{\text{h}}^{(1)}]$ and to spans the

occupied space of the ground state determinant:

$$\langle \phi_{\text{h}}^{(2)} | \phi_{\text{h}}^{(1)} \rangle = 0, \quad (2.12)$$

$$\hat{Q}^{(0)} \phi_{\text{h}}^{(2)} = 0, \quad (2.13)$$

then the determinant $\Phi^{(2)}$ is guaranteed to be orthogonal to $\Phi^{(0)}$ and $\Phi^{(1)}$. The conditions Eqs. (2.12) and (2.13) are sufficient but not necessary to guarantee orthogonality among the first three electronic states.

In the following we describe our constrained multiple hole/particle (CMHP) algorithm, which generalizes Eqs. (2.12) and (2.13) to the case of n electronic states. Due to the complexity of this algorithm, we recommend the reader to follow its description with the help of Fig. 2.3. Suppose we are interested in the excited states that result from the excitation of a given number of core orbitals (n_{c}) and unoccupied orbitals (n_{u}). For convenience we will label the excited state Kohn–Sham determinants $[\Phi^{(i,a)}]$ with two indices, i and a , which stand respectively for the core and unoccupied orbital that are involved in an excited state. Our algorithm starts with a ground-state DFT computation, which yields the determinant $\Phi^{(0,0)}$. Next, we perform a sweep of n_{u} OCDFT computations, which produces the series of solutions:

$$\Phi^{(1,1)}, \Phi^{(1,2)}, \dots, \Phi^{(1,n_{\text{u}})}. \quad (2.14)$$

These solutions are characterized by hole (particle) orbitals that span the occupied (unoccupied) space of $\Phi^{(0,0)}$:

$$\begin{cases} \hat{Q}^{(0)} \phi_{\text{h}}^{(1,a)} = 0 \\ \hat{P}^{(0)} \phi_{\text{p}}^{(1,a)} = 0 \end{cases}, \quad \forall a \leq n_{\text{u}}, \quad (2.15)$$

and particle orbitals that form an orthogonal set:

$$\langle \phi_{\text{p}}^{(1,a)} | \phi_{\text{p}}^{(1,b)} \rangle = \delta_{ab}, \quad \forall a, b \leq n_{\text{u}}. \quad (2.16)$$

In the following iteration of the CMHP algorithm we increase the core index by one and sweep again through a series of solutions:

$$\Phi^{(2,1)}, \Phi^{(2,2)}, \dots, \Phi^{(2,n_{\text{u}})}, \quad (2.17)$$

such that the hole and particle orbitals span respectively the virtual and occupied space of the ground-state determinant:

$$\begin{cases} \hat{Q}^{(0)}\phi_h^{(2,a)} = 0 \\ \hat{P}^{(0)}\phi_p^{(2,a)} = 0 \end{cases}, \quad \forall a \leq n_u, \quad (2.18)$$

and the particle orbitals are orthogonal:

$$\langle \phi_p^{(2,a)} | \phi_p^{(2,b)} \rangle = \delta_{ab}, \quad \forall a, b \leq n_u. \quad (2.19)$$

In addition, we impose orthogonality between the first core orbital of the first sweep, $\phi_h^{(1,1)}$, and the second core orbital of each of the states determined during the current sweep, $\phi_h^{(2,a)}$:

$$\langle \phi_h^{(1,1)} | \phi_h^{(2,a)} \rangle = 0, \quad \forall a \leq n_u. \quad (2.20)$$

This condition guarantees that the hole orbital for the second series of computations is different from the first one. The CMHP algorithm proceeds in a similar way until $i = n_c$. After each sweep over the particle orbitals, a new set of orthogonality conditions is imposed onto the hole orbitals. We notice that the CMHP algorithm does *preserve strict orthogonality among all states*, but it is not equivalent to imposing the strictly minimum orthogonality conditions.

In summary, the CMHP algorithm simplifies the optimization of a series of mutually orthogonal determinants by imposing separate orthogonality conditions onto the core (hole) and valence (particle) orbitals. It is instructive to compare the computational cost of OCDFT and TDDFT. The cost to compute d excited state in OCDFT is proportional to dN^3 and dN^4 for pure and hybrid functional, respectively, since for each excited states one has to solve a set of constrained Kohn–Sham equations. In the case of TDDFT, the computational cost is dominated by the solution of a pseudo-eigenvalue problem, which scales as dN^4 . TDDFT certainly provides a more direct approach to the computation of multiple excited states, but in our experience we observe that the OCDFT/CMHP algorithm has a cost comparable to that of TDDFT and in most cases can be applied in a black-box way.

2.3 Computational Details

Our OCDFT for core-excited states was implemented as a plugin for the PSI4 *ab initio* quantum chemistry package.⁴⁹ The OCDFT excitation energies for a set of 13 small molecules reported in this study were computed using the B3LYP,^{50–53} PBE0⁵⁴, and BLYP^{53,55} functionals, using the correlation-consistent polarized core-valence basis sets (cc-pCVXZ, $X = T, Q$)⁵⁶ and the Karlsruhe valence polarized basis sets (def2-XZVP, $X = T, Q$).^{57,58} All geometries were optimized at the same level of theory as the given excitation energy. The optimized geometries and the NEXAS spectrum of adenine and thymine were computed at the def2-TZVP/B3LYP level of theory (cartesian geometries are provided in the supplementary material). Benchmark TDDFT excitation energies were computed at the B3LYP/def2-QZVP level of theory using the ORCA software package.⁵⁹ All excitation energies reported in this work, including TDDFT, are for singlet states. In the case of OCDFT, the singlet excitation energies were computed via the spin-raising approach described in Ref. 30.

It is mandatory to consider relativistic effects when studying excitations of core electrons.^{60–63} For excitations involving 1s orbitals, we approximate the relativistic excitation energy (ω^R) as the sum of the nonrelativistic excitation energy (ω^{NR}) minus a correction $\Delta\epsilon_{1s}$:

$$\omega^R = \omega^{NR} - \Delta\epsilon_{1s}, \quad (2.21)$$

where $\Delta\epsilon_{1s}$ is the energy difference between the ground-state nonrelativistic (NR) and relativistic (R) Kohn–Sham energies of the 1s orbital:

$$\Delta\epsilon_{1s} = \epsilon_{1s}^R - \epsilon_{1s}^{NR}. \quad (2.22)$$

In this work, the excitation energies computed in OCDFT and TDDFT utilize relativistic orbital energies calculated with first-order Douglas–Kroll–Hess (DKH) Hamiltonian.^{64–66} Relativistic corrections for the 1s orbitals of the second row nuclei range

from 3.8 eV (Si) to 10.1 eV (Cl). In the case of first row 1s core orbitals, $\Delta\epsilon_{1s}$ is negligible (C, N, and O 1s corrections are about 0.1, 0.2, and 0.3 eV, respectively). Similarly, excitations from 2p orbitals of second row elements are negligible (max 0.05 eV) and were not applied to the final results.

The treatment of core-excited states in molecules with symmetry equivalent atoms becomes problematic for both pure and hybrid functionals due to the approximate treatment of exchange and correlation which introduces a self-interaction error.^{67,68} In this case, the symmetry restricted solution produces core holes distributed evenly amongst the symmetry equivalent atoms. Instead, the symmetry unrestricted solution may consist of core holes localized on each individual atom. For all molecules with symmetry equivalent atoms (N_2 , C_2H_2 , and Cl_2) we studied both the symmetry restricted and unrestricted solutions. To obtain a state where the core hole is localized, we utilize a wave function with broken spatial and spin symmetry mixing the coefficients of the alpha and beta orbitals.

Peak intensities for the transition $\Psi^{(n)} \leftarrow \Psi^{(0)}$ are based on the oscillator strength (f_{osc}):

$$f_{osc} = \frac{2}{3} |\mu_{n0}|^2 \omega_n, \quad (2.23)$$

which is calculated from the excitation energies (ω_n) and transition dipole moments (μ_{n0}). Although OCDFT does not provide a direct way to compute transition dipole moments, these can be approximated using the Kohn–Sham determinants as:

$$\mu_{n0} = \langle \Phi^{(n)} | \hat{\mathbf{r}} | \Phi^{(0)} \rangle, \quad (2.24)$$

where $\Phi^{(n)}$ is a generic excited state and $\hat{\mathbf{r}}$ is the position vector. Eq. (2.23) yields the absolute oscillator strength (f_{abs}) for a given transition. The intensity of the spectrum is then scaled relative to the most intense peak, we will refer to these scaled values as the *relative oscillator strength* (f_{rel}). Natural spectroscopic broadening effects are

simulated by convoluting the OCDFT peaks with a Gaussian function whose exponent was calibrated to best fit the experimental spectrum.

2.4 Results and Discussion

2.4.1 Calibration of OCDFT core-excitation energies

The accuracy of OCDFT was benchmarked using a test set that comprises molecules containing first-row (CO, H₂CO, N₂O, N₂, HCN, CH₄, C₂H₄) and second-row elements (SiH₄, PH₃, H₂S, SO₂, HCl, Cl₂), and a total of 40 excited states. Table 2.1 summarizes the performance of OCDFT by reporting the mean absolute error (MAE) in the excitation energy for 12 unique combinations of basis set and density functionals. Individual excitation energies computed with OCDFT and TDDFT are compared to values from gas-phase NEXAS experiments^{69–83} in Tables 3.3 and 3.4. Table 3.3 displays excitations from first-row elements, while Table 3.4 shows excitations from second-row elements. For molecules with symmetry-equivalent atoms (N₂, C₂H₂, and Cl₂), we examined both the symmetrical and symmetry-broken solutions and found that the former have a MAE of 13.7 eV, while the latter yield a MAE of 0.9 eV. Thus all results for highly symmetric molecules reported in Tables 3.3 and 3.4 are based on broken symmetry solutions. The three functionals considered in Table 2.1 produce MAEs ranging from 0.9–1.7 eV. Interestingly, there is no dramatic difference in the accuracy of OCDFT regardless of the amount of Hartree–Fock (HF)

Table 2.1: OCDFT core-excitation energies for a benchmark set composed of 13 diatomic molecules. Mean absolute error (in eV) computed using various combinations of basis sets and density functionals. These statistics refer to a subset of the benchmark set comprised of 35 core-excited states.

Basis Set	Mean Absolute Error (eV)		
	BLYP	B3LYP	PBE0
def2-TZVP	1.0	1.0	0.9
def2-QZVP	1.0	1.0	1.4
cc-pCVTZ	1.3	1.3	1.6
cc-pCVQZ	1.5	1.5	1.7

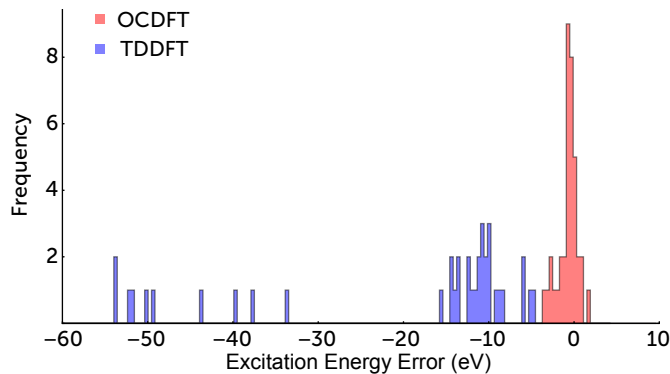


Figure 2.4: Histogram showing the distribution of the error in the computed core excited states. All calculations were done using the B3LYP functional and the def2-QZVP basis set. The red filled bar is OCDFT while the empty blue bar is TDDFT. Notice that the two distributions do not overlap.

exchange present in the functional. Even the BLYP functional, which contains no HF exchange, yields a MAE (1.0 eV using the def2-QZVP basis set) comparable to that of the MAEs produced by its hybrid counterparts. The Karlsruhe family of basis sets yields results that are in better agreement with the experimental excitation energies than the correlation-consistent basis sets.

The average error of OCDFT is commensurate to that of wave function methods for core-excited states. For example, Asmuruf and Besley⁹ reported an average error of 1.2 eV for SOS-CIS(D) applied to a set of excitations similar to the ones used in the present study. While Coriani et al.⁸ reported absolute errors of less than 0.9 eV when applying coupled cluster response theory to a set of carbon, nitrogen, and neon core excitations.

A full comparison of the accuracy of OCDFT and TDDFT core-excitations computed at the B3LYP/def2-QZVP level of theory is shown in Fig. 2.4. The contrast between the two error distributions is striking. As expected, TDDFT performs rather poorly underestimating the excitation energies, on average, by 15 eV and a maximum error of -53.6 eV. On the contrary, OCDFT yields an error distribution peaked near zero and a maximum error of -3.7 eV. The TDDFT error distribution has a peculiar shape, displaying two distinct groups of excited states. The first is a narrow distri-

bution that exists in the range -4 eV to -15 eV, while the second one is broader and ranges from -54 eV to -38 eV. An analysis of the group of excited states with the largest errors reveals that these consist solely of $1s$ core-excitations of second-row elements. This finding is in agreement with previous studies by Nakata⁸⁴ and Besley.²⁴ Since our excitation energies are corrected for relativistic effects (albeit with a crude approximation), the bulk of the error observed for second-row elements must be attributed to a deficiency of the exchange-correlation functional.⁸⁵ This dramatic difference in the performance of TDDFT suggests that it is helpful to separately analyze first-row and second-row core excitations to highlight their distinctive features.

Table 3.3 reports valence and Rydberg excitation energies of $1s$ orbital localized on first row elements (C, N, or O). Compared to experiment, TDDFT produces a mean absolute error of 11.6 eV, and as discussed earlier, this result is consistent with previous studies that used TDDFT with conventional hybrid functionals.¹⁵ OCDFT calculations yield a mean absolute error of 0.4 eV. To put this error into perspective, it can be compared to the performance of TDDFT with the BH^{0.58}LYP functional,²⁴ a reparameterization of the BLYP functional^{50,86} that has been augmented to include

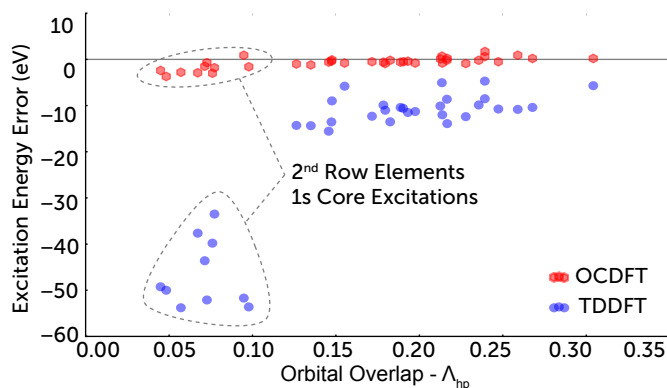


Figure 2.5: Scatterplot displaying the excitation energy error as a function of orbital overlap for a benchmarking test set of 40 core excitations from 13 different molecules. Excitation energies were calculated using the B3LYP functional and def2-QZVP basis set. The overlap integrals were computed with numerical grid integration making use of the Gaussian cube files produced by in Psi4 OCDFT calculations. Grids were calculated with a double zeta basis set and 0.1 grid spacing.

58% HF exchange, 39% B88 exchange, and 8% Slater exchange. When applied to a test set of first row core excitations similar to those in Figure 3.3, BH^{0.58}LYP yielded a mean average error of 0.8 eV. It is encouraging that OCDFT can achieve a higher level of accuracy without altering the amount of Hartree–Fock exchange included in the functional. It is also gratifying to see that the OCDFT MAE for this set of first row core excited states is comparable to the MAE obtained for valence excited states (0.3 eV reported in Ref. 30). When computing core excited states of second-row nuclei, TDDFT becomes highly inaccurate, producing an average error larger than 30 eV. Previous work by Tozer and coworkers²³ showed that there is a correlation between the level of accuracy of TDDFT excitation energies and the amount of overlap between the orbitals involved. We expect this correlation to also be observed in core electron excitations, where the core hole and valence particle orbitals have little overlap. Following Tozer et al.,²³ in OCDFT we define the overlap between the hole and particle orbital (Λ_{hp}) for any excited state n as the integral:

$$\Lambda_{\text{hp}} = \int |\phi_{\text{h}}^{(n)}(\mathbf{r})| |\phi_{\text{p}}^{(n)}(\mathbf{r})| \, \text{d}\mathbf{r}. \quad (2.25)$$

Figure 2.5 reports the distribution of OCDFT and TDDFT excited states as a function of the energy error and the hole/particle orbital overlap. The scatterplot clearly shows that OCDFT is less sensitive to variations in the overlap. When calculating core excited states with $\Lambda_{\text{hp}} < 0.12$, the MAE for OCDFT increases by only 1.5 eV, while in the case of TDDFT the absolute error increases drastically by 35.9 eV.

Tables 3.3 and 3.4 also report oscillator strengths computed with OCDFT at the B3LYP level of theory. An extensive quantitative comparison with experimental line intensities is not practical, however, some qualitative analysis is possible and provides insight into the reliability of the computed oscillator strengths. In general, the lower energy core-valence excitations are more intense than the higher energy Rydberg excitations. For example, an analysis of the experimental K-edge spectrum of carbon monoxide obtained by Domke et al.⁷⁷ shows that the C 1s $\rightarrow \pi^*$ transition

Table 2.2: Core excitation energies for molecules containing first-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 69– 77.

Molecule	Excitation	Exp. (eV)	Error (eV)		$f_{\text{abs}} \times 10^{-2}$
			TDDFT	OCDFT	OCDFT
CO	C 1s \rightarrow π^*	287.4	-11.3	-0.8	4.23
	C 1s \rightarrow 3s	292.4	-10.5	0.9	0.26
	O 1s \rightarrow π^*	534.2	-13.4	-1.2	1.74
H ₂ CO	O 1s \rightarrow 3s	538.9	-13.0	0.2	0.01
	C 1s \rightarrow π^*	286.0	-10.7	-0.6	3.33
	C 1s \rightarrow 3s	290.2	-10.7	-0.2	0.62
N ₂ O [†]	O 1s \rightarrow 3s	535.4	-14.1	-0.6	0.06
	O 1s \rightarrow π^*	530.8	-14.0	-0.8	2.04
	O 1s \rightarrow π^*	534.8	-14.3	-1.0	1.08
N ₂	O 1s \rightarrow 3s	536.7	-13.6	-0.4	1.67
	N _c 1s \rightarrow 3s	407.5	-12.1	0.6	3.11
	N _t 1s \rightarrow π^*	401.1	-12.2	-0.9	2.41
HCN	N _t 1s \rightarrow 3s	404.0	-11.5	-0.4	0.90
	N 1s \rightarrow π^*	401.0	-12.4	-0.9	2.90
	N 1s \rightarrow 3s	406.2	-8.5	1.7	0.00
CH ₄	C 1s \rightarrow π^*	286.4	-10.6	-0.5	2.57
	C 1s \rightarrow 3s	289.1	-9.9	-0.1	0.81
	N 1s \rightarrow π^*	399.7	-12.0	-0.8	2.46
C ₂ H ₂	N 1s \rightarrow 3s	401.8	-10.4	0.2	0.28
	C 1s \rightarrow 3p	288.0	-10.1	0.1	1.85
	C 1s \rightarrow 3s	287.1	-10.8	-0.5	0.00
MAE	C 1s \rightarrow π^*	285.8	-10.5	-0.6	2.27
	C 1s \rightarrow 3s	287.7	-9.1	-0.1	0.11
MAE			11.6	0.4	

[†]The subscripts c and t stand for the center and tail nitrogen of N₂O.

is a sharp, very intense peak at 287.4 eV. While the C 1s \rightarrow 3s transition is a peak of significantly lower intensity at 292.4 eV. OCDFT produces an oscillator strength for the C 1s \rightarrow π^* transition that is an order of magnitude greater than that of the C 1s \rightarrow 3s transition, which is consistent with the observed experimental trend. The same comparison can be done with the oxygen K-edge of carbon monoxide and OCDFT shows similar agreement with the experimental spectrum.

Table 2.3: Core excitation energies for molecules containing second-row elements. Computations were performed using the B3LYP density functional and def2-QZVP basis set. The OCDFT and TDDFT results are reported here as deviations from the experimental value in electron volts (eV), mean absolute error (MAE) is also reported for each method. Experimental values are from Refs. 78–83.

Molecule	Excitation	Exp. (eV)	Error (eV)		$f_{\text{abs}} \times 10^{-2}$
			TDDFT	OCDFT	OCDFT
SiH ₄	Si 1s \rightarrow σ^*	1842.5	-38.4	-1.8	0.18
	Si 2p \rightarrow σ^*	102.8	-4.8	0.6	0.33
PH ₃	P 1s \rightarrow σ^*	2145.8	-44.1	-2.9	0.23
	P 2p \rightarrow σ^*	132.3	-5.1	0.7	0.36
H ₂ S	S 1s \rightarrow σ^*	2473.1	-48.3	-3.0	0.14
	S 1s \rightarrow 4p	2476.3	-52.1	-1.5	0.24
	S 2p \rightarrow σ^*	164.5	-5.1	0.8	1.98
	S 2p \rightarrow 4s	166.5	-7.1	-0.7	0.51
SO ₂	S 1s \rightarrow π^*	2473.8	-50.1	-3.7	0.39
	S 1s \rightarrow 4p	2478.4	-49.3	-2.4	0.24
	S 2p \rightarrow 4s	171.3	-8.3	-1.5	0.05
HCl	Cl 1s \rightarrow σ^*	2823.9	-53.8	-2.3	0.21
	Cl 1s \rightarrow 4p	2827.8	-52.1	-0.7	0.22
	Cl 2p \rightarrow σ^*	201.0	-6.1	0.8	0.00
Cl ₂	Cl 1s \rightarrow σ^*	2821.3	-53.6	-1.6	0.35
	Cl 1s \rightarrow 4p	2828.5	-51.7	0.9	0.14
	Cl 2p \rightarrow σ^*	198.7	-5.7	-0.8	0.00
MAE			31.6	1.6	

To understand why OCDFT outperforms TDDFT, we will consider a model consisting of two electrons in two orbitals (ϕ_h , ϕ_p) of different symmetry.^{30,38,91} This model makes it possible to compare the TDDFT and OCDFT excitation energies to that of CIS. As illustrated in Fig. 2.5, core excitations are characterized by a very small overlap between the hole and particle orbital. Therefore, our analysis considers the limit of zero overlap ($\Lambda_{\text{hp}} = 0$). For a functional containing a given fraction (a) of Hartree–Fock exchange, the TDDFT and CIS singlet excitation energies (ω_s) for our model differ approximately by:

$$\omega_s^{\text{TDDFT}} - \omega_s^{\text{CIS}} \cong (1 - a)[v_p^x - v_h^x] + (1 - a)[J_{\text{ph}} - J_{\text{hh}}], \quad (2.26)$$

where $v_i^x = (\phi_i | v_x | \phi_i)$ is an exchange potential integral and J_{ij} is the Coulomb repul-

sion integral $(\phi_i\phi_i|r_{12}^{-1}|\phi_j\phi_j)$. When $a \ll 1$, the right-hand side of Eq. 2.26 contains three local integrals v_h^x , v_p^x , and J_{hh} . However, the Coulomb repulsion integral between the hole and particle orbitals (J_{ph}) is nonlocal and causes TDDFT to incorrectly describe the physics of the hole/particle pair. On the contrary, when $a = 1$ there is exact cancellation of the nonlinear terms and TDDFT is equivalent to CIS. This is consistent with the observation that increasing the amount of Hartree–Fock exchange improves the description of core-excited states.^{26,87–90}

The same analysis finds that the OCDFT and CIS excitation energies differs by a sum of local self-interaction terms:

$$\begin{aligned} \omega_s^{\text{OCDFT}} - \omega_s^{\text{CIS}} \cong & (1 - a)[v_p^x - v_h^x + \frac{1}{2}J_{pp} - \frac{1}{2}J_{hh} \\ & + \frac{1}{2}(\text{hh}|\hat{f}_x|\text{hh}) + \frac{1}{2}(\text{pp}|\hat{f}_x|\text{pp})], \end{aligned} \quad (2.27)$$

where $(ii|\hat{f}_x|ii)$ is an exchange kernel integral. As observed by Ziegler and co-workers in the case of charge-transfer excitations computed via the constricted variational DFT,³⁸ there is partial cancellation between the local integrals that appear in Eq. (2.27) and the excitation energy has the correct asymptotic behavior. Our observation that the amount of Hartree–Fock exchange has little influence on the computed core excitation energies, suggests that even in the case of OCDFT there is partial cancellation of the terms in Eq. (2.27) that appear in square brackets. Thus, although charge-transfer and core-excited states are very different in nature, this simple model suggests a formal connection between the two.

2.4.2 Application to Nucleobases: Thymine and Adenine Near-Edge Spectra

Nucleobases play a key biological role as the building blocks of DNA and recently show promise as potential materials for electronic/technological applications.^{92–95} Early X-ray studies of nucleobases consisted of scattering and diffraction experiments.^{96–99} The first near-edge absorption experiments were performed in the 90s by Mitra-Kirtley et

al.¹⁰⁰ These authors specifically targeted the nitrogen core electrons and probed the sensitivity of the $1s \rightarrow \pi^*$ resonances to the surrounding chemical environment. More recent experiments have moved beyond simple characterization of the intramolecular environment and aim to probe intermolecular interactions of nucleobases with metal surfaces.^{101–104} At the same time a wide array of computational methods have been used to compute the near-edge structure of nucleobases, including: restricted active space SCF (RASSCF),¹⁰⁵ improved virtual orbital SCF (IVO-SCF)¹⁰⁶, a complex polarization propagator method (CPP)¹⁰⁷, SIC-DFT,¹⁰⁸ a DFT transition potential method (DFTTP),¹⁰⁶ the equivalent core approximation (ECA) method,¹⁰⁹ and the second-order algebraic diagrammatic construction [ADC(2)] method.^{110,111}

Here we present an OCDFT simulation of the gas-phase NEXAS spectrum of thymine and adenine. We first give an overview of the performance of OCDFT relative to the gas-phase NEXAS experiments done by Plekan et al.¹¹⁰ This is followed by an in-depth analysis of the spectral features simulated in OCDFT. Last we compare our work to previous studies employing ADC(2) theory. The numbering schemes used for adenine and thymine are shown in Tables 2.4 and 2.6. We follow a widely used convention of numbering the atoms according to the Hartree–Fock core orbital energy. The relevant ground-state virtual orbitals for thymine and adenine are shown in Figs. 2.6 and 2.8, respectively. The simulated and experimental NEXAS spectra of thymine and adenine are reported in Figs. 2.7 and 2.9.

Overall Performance

The simulated thymine and adenine NEXAS spectra shown in Figs. 2.7 and 2.9 agree well with the experimental data. Tables 2.4 and 2.6 report the dominant contributions to the NEXAS spectra (full tables showing all contributions are provided in the supplementary material), along with excitation energies and relative oscillator strengths. We also report the nature of each excited state. That is, for each transition we specify the core $1s$ electron excited (X_i where $X = O, N, C$ and i is the label of the atom

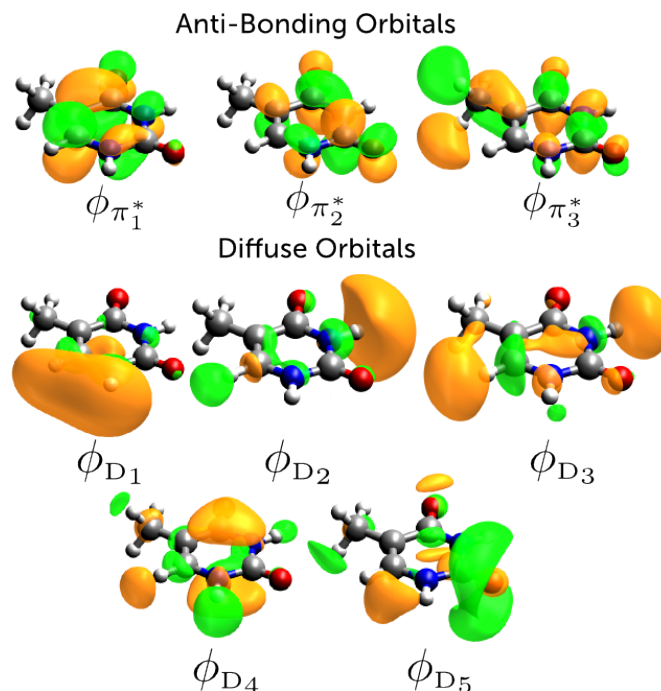


Figure 2.6: Relevant virtual orbitals for thymine numbered in according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is diffused are labeled as D.

in our numbering scheme), and the ground-state virtual orbital that has the most overlap with the particle orbital, together with its weight. The experimental energies reported are the peak maxima for each spectral feature, and can be approximated by the OCDFT transition in that region with the strongest oscillator strength. When using peak maxima as a comparison, OCDFT represents the thymine spectrum with an average error of 0.3 eV, and that of adenine with an average error of 0.1 eV. A common feature of NEXAS spectra is the appearance of multiple low-intensity transitions in the higher energy regions. This is represented well by OCDFT as evidenced by the stick spectrum shown in Figs. 2.7 and 2.9 where the higher energy regions are populated by multiple low intensity transitions. We emphasize that the computed OCDFT spectra are obtained from unshifted excitation energies.

Thymine Oxygen K-Edge

Figure 2.7a displays the experimental and theoretical oxygen K-edge[?] spectra of thymine. The low energy regime of the oxygen K-edge is dominated by two high intensity peaks. Peak A results from the transition $O_2 \rightarrow \phi_{\pi_2^*}$, while peak B results

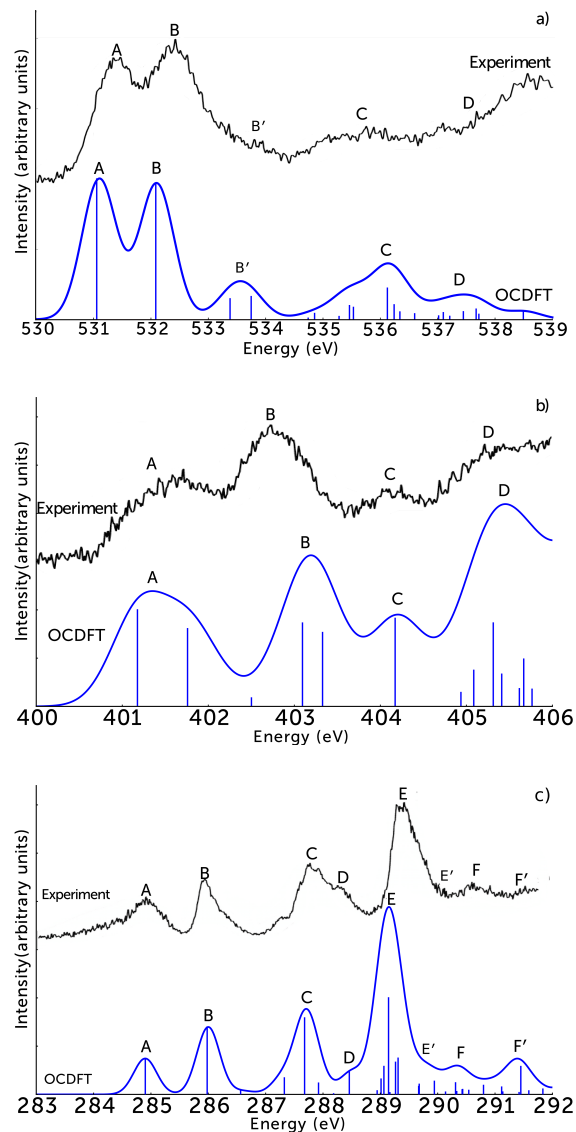


Figure 2.7: Core excited states for thymine computed using B3LYP functional and def2-TZVP basis set. The OCDFT oxygen and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV, while the carbon spectrum uses a Gaussian with FWHM equal to 0.2 eV. The experimental spectra are reproduced with permission from Ref. 110.

from the transition $O_1 \rightarrow \phi_{\pi_1^*}$ with a peak intensity that is roughly equal to that of peak A. Experimentally, A and B are centered at 531.4 eV and 532.3 eV, respectively, and are predicted by OCDFT to within 0.3 eV. The $O_2 \rightarrow \phi_{\pi_2^*}$ transition is predicted as the peak of highest intensity with $f_{\text{abs}} = 0.02$, and this result is consistent with the ADC(2) analysis performed by Plekan et al.¹¹⁰ which predicts $f_{\text{abs}} = 0.03$. According to OCDFT, the shoulder feature B' is composed of $O_1 \rightarrow \phi_{\pi_2^*}$ and $O_2 \rightarrow \pi_1^*$ transitions predicted to have a fairly strong oscillator strength ($f_{\text{rel}} \approx 0.2$), which is in discrepancy with the low intensity peaks observed in the experimental spectrum. As stated earlier, excitations of weaker intensity are characteristic of the higher energy regime of the K-edge and have been attributed to strong mixing of core-valence excited states with Rydberg excited states of similar energy.¹¹² This strong mixing causes excitations to be spread out over several different final states, resulting in transitions of weak intensity. The mixing in this region of the spectrum makes it difficult to classify specific transitions experimentally. OCDFT results show that peak C is largely composed of a mixture of diffuse Rydberg excitations within the energy interval of 534.7 to 536.4 eV. While the majority of the contributions to peak D, are excitations to ϕ_{D_2} and ϕ_{D_5} , with $f_{\text{rel}} < 0.1$, peaks C and D both have $\phi_{\pi_3^*}$ character. In both cases, these resonances are weak and overshadowed by multiple Rydberg transitions in both cases.

Thymine Nitrogen K-Edge

The K-edge pictured in Figure 2.7b is characterized by four distinct spectral peaks. These lowest energy contributions to peak A are excitations from N_3 and N_4 to $\phi_{\pi_2^*}$. OCDFT predicts their excitation energies to be 401.8 eV and 401.2 eV, respectively. The experimental peak maximum is at 401.7 eV, which agrees well with the Gaussian profile shown in the OCDFT spectrum. Rydberg transitions from the N_3 and N_4 to ϕ_{D_1} are the dominant resonances contributing to the character of peak B, along with a valence excitation $N_4 \rightarrow \phi_{\pi_1^*}$ predicted at 402.5 eV. This agrees well with

the experimental peak assignment at 402.7 eV. A very intense $N_4 \rightarrow \phi_{D_3}$ transition accounts for the peak at 404.1 eV. OCDFT simulates this peak perfectly with a Gaussian centered at 404.2 eV. peak D is the amalgamation of two π^* resonances and multiple Rydberg states with the π^* resonances being the transitions of strongest intensity. Excitation energies of these π^* resonances agree well with the experimental peak assignment at 405.5 eV, with $N_4 \rightarrow \phi_{\pi_3^*}$ at 405.3 eV and $N_3 \rightarrow \phi_{\pi_1^*}$ at 405.7 eV.

Thymine Carbon K-Edge

The shape of the carbon K-edge displayed in Figure 2.7c is governed by four strong π^* resonances. Unique to the carbon K-edge is the fact that the strongest transition is not the lowest energy π^* resonance, the $C_5 \rightarrow \phi_{\pi_1^*}$ transition is a relatively high energy excitation and produces the strongest peak intensity, despite close proximity to several Rydberg states. Peak A at 284.9 eV is the result of the transition $C_8 \rightarrow \phi_{\pi_2^*}$, the position of this peak is predicted exactly by OCDFT. A slightly stronger transition at 285.9 eV is mostly due to a $C_7 \rightarrow \phi_{\pi_2^*}$ excitation, with small contribution from another $\phi_{\pi_2^*}$ transition resulting from an excitation from the C_9 core. Peaks C and D have experimental peak maxima at 287.8 eV and 288.4 eV respectively, and blend together to form one band. Both contributions to the spectral band are represented well by OCDFT. Peak E is mostly composed of core excitations to diffuse orbitals, however these transitions have relatively weak intensities compared to the strong $\phi_{\pi_2^*}$ transition predicted at 289.1 eV. Excitations to diffuse orbitals are dominant contributors to the remaining spectral features F and F'.

Adenine Nitrogen K-Edge

Figure 2.9a displays the nitrogen k-edge of adenine computed with OCDFT. The most dominant feature of the spectrum, peak A, is experimentally classified at 399.5 eV, OCDFT predicts that this feature is a result of core-valence excitations originating from three of the four nitrogens on the purine ring system (N_4, N_3, N_5). The largest

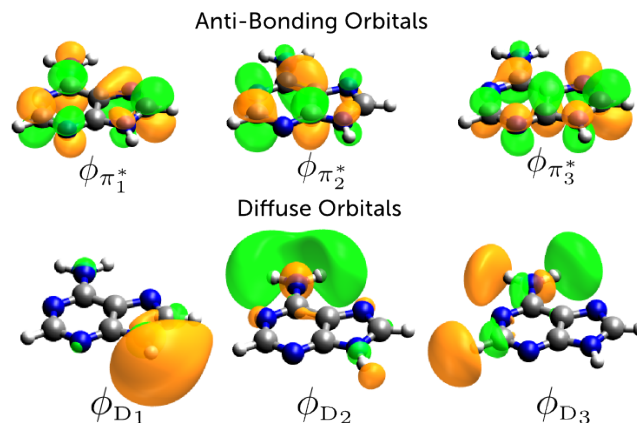


Figure 2.8: Relevant virtual orbitals for adenine numbered according to the orbital energy. Orbitals with obvious π^* character are labeled as such, while orbitals where electron density is mostly diffused on the outside of the molecule are labeled as ϕ_D .

contributor being a transition from $N_5 \rightarrow \phi_{\pi_2^*}$ predicted at 399.4 eV, which agrees well with the experimental classification for the peak. An apparent shoulder feature with fairly weak oscillator strength is present in the experimental spectrum in the region from 399.8 – 400.4 eV. OCDFT represents this feature well, with two weak π^* resonances resulting from the two nitrogens on the six-membered ring. A few rising shoulder features are shown in the experimental spectrum in the region from 401.0 eV – 401.3 eV, OCDFT predicts a $N_3 \rightarrow \phi_{\pi_2^*}$ transition in this region with a relative oscillator strength of 0.094. Peak B is a mixture of transitions to π^* orbitals as well as diffuse orbitals, with the π^* resonances being the prominent contributors. The experimental spectrum shows a relatively weak resonance around 403.0 (peak C). We predict that the dominant contributor to peak C is a transition from $N_2 \rightarrow \phi_{D_3}$, the intensity of which, that rivals the strongest transition (peak A $f_{\text{rel}} = 0.925$). This peak intensity is contrary to the experimental spectrum which shows peak C as a superposition of weak transitions. The highest energy transitions are all weak transitions to mostly orbitals of diffuse character, with transitions getting more intense as they approach 405.0 eV.

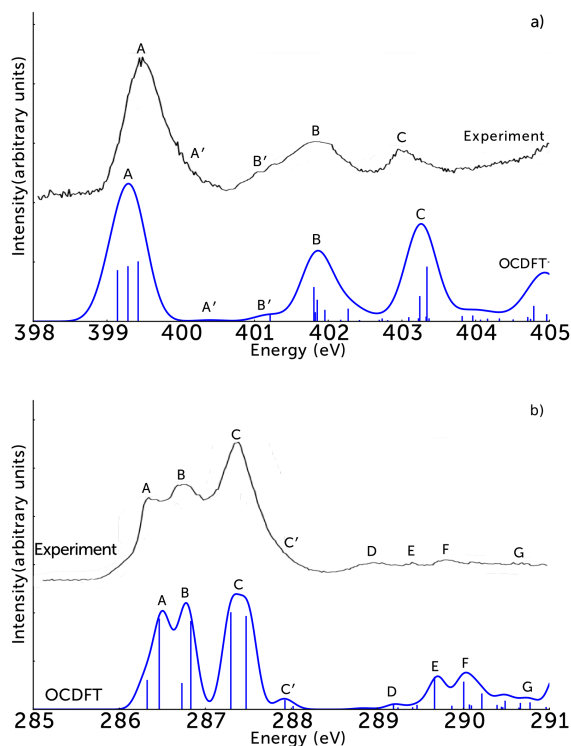


Figure 2.9: Core excited states for adenine computed using B3LYP functional and def2-TZVP basis set. The OCDFT carbon and nitrogen K-edge spectra were convoluted with a Gaussian function with full width at half maximum (FWHM) equal to 0.3 eV. The experimental spectra are reproduced with permission from Ref. 110.

Adenine Carbon K-Edge

Fig. 2.9b compares the OCDFT and experimental carbon K-edge of adenine. The experimental carbon K-edge for adenine is dominated by a large single band with three distinct resonances in the low energy regime. The theoretical spectrum shows peaks A and B, blending together into a single band, the experimental spectrum shows these peaks at similar intensities, with peak B being slightly more intense, this is represented well in our calculated spectrum. Spectral positions for peaks A, B, and C are all in good agreement with experiment, however, the oscillator strength is inconsistent with experimental peak intensities. According to the experimental results, peak C should be roughly 50% more intense than the adjacent peak B. OCDFT predicts that the $C_6 \rightarrow \phi_{\pi^*}$ that dominates peak C is only slightly more intense than the dominant

contributions to peaks A and B. A very weak shoulder feature is present on the falling edge of peak C, we denote this feature as C' and it is represented well by OCDFT. This shoulder results from π^* transitions with hole orbitals located on the bridge carbons (C₈ and C₁₀) and a carbon located on the six-member ring (C₉). Peaks D-G were not assigned due to their low intensities. OCDFT reveals that peak D is a very weak spectral feature with dominant contributions from excitations to diffuse orbital ϕ_{D_3} . Peaks E and F are also weak spectral features resulting mostly from transitions to diffuse orbitals. Every transition with energy higher than peak F is extremely weak

Comparison with ADC(2) Calculations

Previous theoretical studies performed using ADC(2) allow us to assess the accuracy of our adenine and thymine OCDFT spectra. Three key differences in the spectra are noted here. The shoulder feature B' in the thymine oxygen K-Edge shown in Fig. 2.7 is absent from the ADC(2) spectrum.¹¹⁰ However, a more recent study by Wenzel, Wormit, and Dreuw¹¹¹ uses a core-valence separation (CVS) approximation to the ADC(2) working equations [CVS-ADC(2)] and predicts three excitations in this shoulder region B', all with $f_{\text{abs}} < 0.001$. These weaker oscillator strengths predicted by CVS-ADC(2) are more consistent with the experimental peak profile.

The overall shape of the OCDFT thymine nitrogen K-Edge is more consistent with the experimental excitation manifold than the ADC(2) spectrum. The two peaks A and B in Fig. 2.7b have clear, distinct maxima which are produced well quantitatively with ADC(2) (after applying a uniform shift of -2.59 eV to the spectrum), strong π^* resonances are reported near both experimental peak maxima. However, the contour of the peak is inconsistent with the experimental manifold. The ADC(2) spectrum blends into one large spectral band over the interval from 401.0 eV to 404.5 eV encompassing very closely spaced transitions, all with relatively high oscillator strengths. The extremely tight spacings and high intensities of these transitions seem to be

present even in the updated CVS-ADC(2) results. The OCDFT spectrum does not suffer from this single band issue, as the π^* resonances in peak A are well separated from the strong ϕ_{D_1} transitions in peak B by more than 1.0 eV. The agreement of these results with experiment, suggest that well-separated π^* and Rydberg resonances are more congruous with reality. However, a more detailed study of the nitrogen core excitation manifold of thymine is required to verify this observation. ADC(2) was unable to fully resolve peaks B', B, and C in the adenine nitrogen edge shown in Fig. 2.9a. On the contrary, the OCDFT spectrum represents these peaks well, as separated spectral features, in compliance with the experimental result.

Table 2.5 shows a direct comparison of the OCDFT peak assignment with those obtained from ADC(2). The comparison is restricted to the transitions of highest intensity at each K-edge. All peak assignment are consistent between the two methods and we observe only small deviations in the computed excitation energy (max error 0.54 eV). The excellent agreement of OCDFT and ADC(2) for the highest intensity transitions is encouraging, and suggests that OCDFT could be a very useful tool to aid the assignment of NEXAS spectra.

2.5 Conclusions

In this work we have extended OCDFT in order to calculate multiple core-excited states of first- and second-row elements. We present two developments in OCDFT theory: 1) we show that core excitations can be easily targeted by selecting *hole* orbitals that correspond to core electrons, and 2) we proposed a generalized set of orthogonality conditions that can be used to compute multiple excited states. Our benchmark computations on core excitations from 1s and 2p orbitals of first- and second-row elements using conventional pure and hybrid functionals yield excitation energies with a mean absolute error (MAE) of 1.0 eV. OCDFT excitation energies are slightly more accurate for first-row elements (MAE = 0.4 eV) than second-row

elements (MAE = 1.6 eV).

There are a few potential sources of the remaining error in the OCDFT excitation energies. First, is the fact that the (in principle) exact OCDFT excited state functional is approximated with the ground-state exchange-correlation functional. Second, in our comparison with experiment we use vertical excitation energies, which neglect vibrational effects. Third, our treatment of relativistic effects is rather simplistic. It can only approximately account for the relaxation of the core-hole and neglects relaxation effects of valence orbitals. Last, OCDFT cannot properly handle states that have a multideterminantal character. We expect that the use of an approximate excited state functional is the most important source of error, and that the relativistic and configuration mixing errors are perhaps the least relevant in the case of first- and second-row elements.

The results of our formal analysis and benchmark computations suggests that there are two factors contributing to the success of OCDFT. As in the case of other variational and constrained DFT approaches, OCDFT can properly describe charge-transfer excitations. In addition, because the wave function of the auxiliary system is only subject to orthogonality constraints, OCDFT can fully describe orbital relaxation. To demonstrate this, we consider the NEXAS spectrum of CO₂, which was suggested as a test case by one of the reviewers. Previous computational studies of CO₂¹¹³ showed that B3LYP/TDDFT underestimates the oxygen K-edge by 14 eV. MRCI/CIS and MRCI/CISD underestimate the same feature by 8 and 4.2 eV, respectively. In the case of MRCI/CIS and MRCI/CISD, the residual error was attributed to the lack of orbital relaxation effects. A broken-symmetry OCDFT computation predicts the dominant feature of the oxygen K-edge to be at 534.4 eV, which corresponds to a deviation from experiment of 0.6 eV. This result suggests that relaxation effects do indeed play an important role in the oxygen K-edge of CO₂ and that OCDFT can fully account for them.

Excitation energies computed with TDDFT are significantly less accurate and degrade as one goes from first-row (MAE = 11.6 eV) to second-row elements (MAE = 31.6 eV). Moreover, we show that in OCDFT the choice of the functional and the amount of Hartree–Fock exchange has little effect on the accuracy of the computed core excitation energies. We perform a formal comparison of OCDFT, TDDFT, and CIS excitation energies for an excitation in which the core and virtual orbitals have zero overlap. This analysis suggests that OCDFT’s superior performance can be ascribed to the local nature of the integrals that appear in the expression for the excitation energy.

Our gas-phase OCDFT X-ray absorption spectra of thymine and adenine are in excellent agreement with experiments. OCDFT reproduces all the characteristic features of the NEXAS spectra of these molecules,^{110,111} including the distinct π^* transitions in the lower energy regime and the significant mixing between the π^* and diffuse orbitals in the higher energy regime of the spectra. In addition, OCDFT assignments of the spectral features are in excellent agreement with those made using the ADC(2) method.¹¹⁰ This study shows that our OCDFT approach for core-excited states is a practical and useful tool for the interpretation of NEXAS experiments. From the computational point of view, the scaling of OCDFT vs. the number of electrons (N) is identical to that of ground state DFT (N^3 and N^4 for pure and hybrid functionals) and it is lower than the second-order approximate coupled cluster method (CC2)¹¹⁴ and [ADC(2)], which scale as N^5 .

Two classes of systems that are worth further exploration are transition metal complexes and open-shell molecules. Core excitations in transition metal complexes present a number of additional challenges that will require further extension of the present theory. One of the biggest improvements that will be necessary is a proper treatment of relativistic effects. While the scheme used in this work was sufficient for first- and second-row elements, it will likely prove ineffective for the treatment of

transition metals where both scalar and spin-orbit relativistic effects play an important role. Scalar relativistic effects can be accounted for by combining OCDFT with spin-free approximate relativistic Hamiltonians. One of our immediate goals is to combine OCDFT with the one-electron spin-free version of the exact two-component approach.^{115–120} This improvement will provide a more consistent way to introduce scalar relativistic effects, and will be essential to compute accurate K-edge spectra of elements past the second row. The simulation of L-edge spectra presents additional challenges¹¹ due to the strong mixing of excitations from degenerate 2p core orbitals and the necessity to account for the coupling of molecular multiplets that experience strong spin-orbit coupling. In this respect, the current formulation of OCDFT—which is ideal for excitations that are dominated by a single Slater determinant—cannot properly treat multideterminantal electronic states that arise in L-edge excitations. One way to overcome this limitation is to employ the basis of non-orthogonal determinants that are generated in an OCDFT computation in a subsequent configuration interaction procedure, as described in the original formulation of OCDFT. This solution will certainly present some challenges, but if successful, could be used to account for the coupling of various molecular multiplets via spin-orbit interactions. These considerations also apply to species with a high-spin open-shell ground state and whose excited states cannot be represented by a single Slater determinant. A case that is more problematic is that of molecules with a low-spin open-shell ground state. In this situation, if the ground state DFT calculation yields an unphysical result, then it will be unlikely for OCDFT to yield accurate excitation energies.

2.6 Acknowledgements

W.D.D is currently supported by the National Science Foundation Graduate Research Fellowship Program. W.D.D would also like to acknowledge initial funding from the Emory Initiative for Maximizing Student Development of the National Institutes of

Health under award number R25GM099644. We would also like to acknowledge start-up funds from Emory University. The content presented in this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation or the National Institute of Health.

Appendix A: OCDFT Equations for the Constrained Multiple Hole/Particle Approach

In this appendix we report details of the algorithm used to compute multiple solutions of the OCDFT equations via the constrained multiple hole/particle method. The OCDFT equations are solved following the sequence:

$$\begin{aligned}
 (i, a) : (0, 0) &\rightarrow (1, 1) \rightarrow (1, 2) \rightarrow \cdots \rightarrow (1, n_u) \rightarrow \\
 &\rightarrow (2, 1) \rightarrow (2, 2) \rightarrow \cdots \rightarrow (2, n_u) \rightarrow \\
 &\vdots \\
 &\rightarrow (n_c, 1) \rightarrow (n_c, 2) \rightarrow \cdots \rightarrow (n_c, n_u),
 \end{aligned} \tag{2.28}$$

where n_c and n_u are the number of core and unoccupied orbitals, respectively, and $n_c n_u$ is the total number of excited states computed.

The OCDFT equations consist of a set of three coupled eigenvalue equations:

$$\begin{aligned}
 (1 - \hat{P}_{\text{h/p}}^{(i,a)}) \hat{f}^{(i,a)} (1 - \hat{P}_{\text{h/p}}^{(i,a)}) |\phi_k^{(i,a)}\rangle &= \epsilon_k^{(i,a)} |\phi_k^{(i,a)}\rangle, \\
 \hat{P}^{(0)} (1 - \hat{Q}_s^{(i,a)}) \hat{f}^{(i,a)} (1 - \hat{Q}_s^{(i,a)}) \hat{P}^{(0)} |\phi_h^{(i,a)}\rangle &= \epsilon_h^{(i,a)} |\phi_h^{(i,a)}\rangle, \\
 \hat{Q}^{(0)} (1 - \hat{P}_s^{(i,a)}) \hat{f}^{(i,a)} (1 - \hat{P}_s^{(i,a)}) \hat{Q}^{(0)} |\phi_p^{(i,a)}\rangle &= \epsilon_p^{(i,a)} |\phi_p^{(i,a)}\rangle,
 \end{aligned} \tag{2.29}$$

where $\hat{f}^{(i,a)}$ is the Kohn–Sham Hamiltonian operator computed using the density corresponding to the state $\Phi^{(i,a)}$. The projection operators that enter the OCDFT

equations are defines as:

$$\hat{P}_{h/p}^{(i,a)} = \hat{P}_h^{(i,a)} + \hat{P}_p^{(i,a)} \quad (2.30)$$

$$\hat{P}_h^{(i,a)} = \sum_{j < i}^{\text{holes}} |\phi_h^{(j,1)}\rangle \langle \phi_h^{(j,1)}| \quad (2.31)$$

$$\hat{P}_p^{(i,a)} = \sum_{b < a}^{\text{particles}} |\phi_p^{(i,b)}\rangle \langle \phi_p^{(i,b)}|, \quad (2.32)$$

$$\hat{P}_s^{(i,a)} = \hat{P}^{(i,a)} - \hat{P}_p^{(i,a)}, \quad (2.33)$$

$$\hat{Q}_s^{(i,a)} = \hat{Q}^{(i,a)} - \hat{P}_h^{(i,a)}. \quad (2.34)$$

Bibliography

- [1] McMillan, E. M. *Phys. Rev.* **1945**, *68*, 143–144.
- [2] Hua, W.; Gao, B.; Li, S.; Ågren, H.; Luo, Y. *J. Phys. Chem. B* **2010**, *114*, 13214–13222.
- [3] Contini, G.; Carravetta, V.; Di Castro, V.; Stranges, S.; Richter, R.; Alagia, M. *J. Phys. Chem. A* **2001**, *105*, 7308–7314.
- [4] Hähner, G. *Chem. Soc. Rev.* **2006**, *35*, 1244–1255.
- [5] Guo, Q. X.; Senda, H.; Saito, K.; Tanaka, T.; Nishio, M.; Ding, J.; Fan, T. X.; Zhang, D.; Wang, X. Q.; Liu, S. T.; Shen, B.; Ohtani, R. *Appl. Phys. Lett.* **2011**, *98*, 181901.
- [6] Feneberg, M.; Röppischer, M.; Esser, N.; Cobet, C.; Neuschl, B.; Meisch, T.; Thonke, K.; Goldhahn, R. *Appl. Phys. Lett.* **2011**, *99*, 021903.
- [7] Nakazato, T.; Oyamada, M.; Niimura, N.; Urasawa, S.; Konno, O.; Kagaya, A.; Kato, R.; Kamiyama, T.; Torizuka, Y.; Nanba, T.; Kondo, Y.; Shibata, Y.; Ishi, K.; Ohsaka, T.; Ikezawa, M. *Phys. Rev. Lett.* **1989**, *63*, 1245–1248.
- [8] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A* **2012**, *85*, 022507.
- [9] Asmuruf, F. A.; Besley, N. A. *Chem. Phys. Lett.* **2008**, *463*, 267–271.
- [10] Roemelt, M.; Neese, F. *J. Phys. Chem. A* **2013**, *117*, 3069–3083.
- [11] Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 204101.
- [12] Schirmer, J. *Phys. Rev. A* **1982**, *26*, 2395.
- [13] Trofimov, A. B.; Schirmer, J. *J. Phys. B: At. Mol. Opt. Phys.* **1995**, *28*, 2299.
- [14] Sheehy, J. A.; Gil, T. J.; Winstead, C. L.; Farren, R. E.; Langhoff, P. W. *J. Chem. Phys.* **1989**, *91*, 1796–1812.
- [15] Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. *J. Chem. Phys.* **2009**, *130*, 124308.
- [16] Ågren, H.; Jensen, H. J. A. *Chem. Phys.* **1993**, *172*, 45–57.

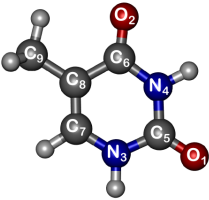
- [17] Triguero, L.; Pettersson, L. G. M.; Ågren, H. *Phys. Rev. B* **1998**, *58*, 8097–8110.
- [18] Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- [19] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.
- [20] Runge, E.; Gross, E. K. *Phys. Rev. Lett.* **1984**, *52*, 997.
- [21] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.
- [22] DeBeer George, S.; Petrenko, T.; Neese, F. *J. Phys. Chem. A* **2008**, *112*, 12936–12943.
- [23] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.
- [24] Besley, N. A.; Peach, M. J. G.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10350.
- [25] Tu, G.; Carravetta, V.; Vahtras, O.; Ågren, H. *J. Chem. Phys.* **2007**, *127*, 174110.
- [26] Nakata, A.; Imamura, Y.; Otsuka, T.; Nakai, H. *J. Chem. Phys.* **2006**, *124*, 094105.
- [27] Capano, G.; Penfold, T.; Besley, N.; Milne, C.; Reinhard, M.; Rittmann-Frank, H.; Glatzel, P.; Abela, R.; Rothlisberger, U.; Chergui, M.; Tavernelli, I. *Chem. Phys. Lett.* **2013**, *580*, 179–184.
- [28] Besley, N. A.; Noble, A. *J. Phys. Chem. C* **2007**, *111*, 3333–3340.
- [29] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024.
- [30] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- [31] Kowalczyk, T.; Yost, S. R.; Van Voorhis, T. *J. Chem. Phys.* **2011**, *134*, 054128.
- [32] Ziegler, T.; Rauk, A.; Baerends, E. J. *Theor. Chim. Acta.* **1977**, *43*, 261–271.
- [33] Wu, Q.; Van Voorhis, T. *J. Chem. Theory Comput.* **2006**, *2*, 765–774.
- [34] Grling, A. *Phys. Rev. A* **1999**, *59*, 3359.
- [35] Ziegler, T.; Seth, M.; Krykunov, M.; Autschbach, J.; Wang, F. *J. Chem. Phys.* **2009**, *130*, 154102.
- [36] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Theory Comput.* **2011**, *7*, 2485–2491.
- [37] Krykunov, M.; Ziegler, T. *J. Chem. Theory Comput.* **2013**, *9*, 2761–2773.
- [38] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Phys.* **2012**, *136*, 124107.
- [39] Baruah, T.; Pederson, M. R. *J. Chem. Theory Comput.* **2009**, *5*, 834–843.
- [40] Olguin, M.; Zope, R. R.; Baruah, T. *J. Chem. Phys.* **2013**, *138*, 074306.
- [41] Zope, R. R.; Olguin, M.; Baruah, T. *J. Chem. Phys.* **2012**, *137*, 084317.
- [42] Theophilou, A. K. *J. Phys. C* **1979**, *12*, 5419.
- [43] Fritsche, L. *Phys. Rev. B* **1986**, *33*, 3976.
- [44] Gross, E. K. U.; Oliveira, L. N.; Kohn, W. *Phys. Rev. A* **1988**, *37*, 2805.

- [45] Gross, E. K. U.; Oliveira, L. N.; Kohn, W. *Phys. Rev. A* **1988**, *37*, 2809.
- [46] Levy, M.; Nagy, A. *Phys. Rev. Lett.* **1999**, *83*, 4361.
- [47] Nagy, A.; Levy, M. *Phys. Rev. A* **2001**, *63*, 052502.
- [48] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A* **2012**, *85*, 042518.
- [49] Turney, J. M. et al. *WIREs Comput Mol Sci* **2012**, *2*, 556–565.
- [50] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- [51] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- [52] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [53] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [54] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
- [55] Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200–206.
- [56] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.
- [57] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [58] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- [59] Neese, F. *WIREs Comput Mol Sci* **2012**, *2*, 73–78.
- [60] Maganas, D.; Roemelt, M.; Weyhermüller, T.; Blume, R.; Hävecker, M.; Knop-Gericke, A.; DeBeer, S.; Schlögl, R.; Neese, F. *Phys. Chem. Chem. Phys.* **2014**, *16*, 264.
- [61] DeBeer George, S.; Neese, F. *Inorg. Chem.* **2010**, *49*, 1849–1853.
- [62] Bauer, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 13827.
- [63] Ankudinov, A. L.; Rehr, J. J.; Low, J. J.; Bare, S. R. *J. Chem. Phys.* **2002**, *116*, 1911.
- [64] Douglas, M.; Kroll, N. M. *Ann. Phys.* **1973**, *82*, 89.
- [65] Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756–763.
- [66] Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742.
- [67] Bally, T.; Sastry, G. N. *J. Phys. Chem. A* **1997**, *101*, 7923–7925.
- [68] Lundberg, M.; Siegbahn, P. E. M. *J. Chem. Phys.* **2005**, *122*, 224103.
- [69] Püttner, R.; Dominguez, I.; Morgan, T. J.; Cisneros, C.; Fink, R. F.; Rotenberg, E.; Warwick, T.; Domke, M.; Kaindl, G.; Schlachter, A. S. *Phys. Rev. A* **1999**, *59*, 3415.
- [70] Remmers, G.; Domke, M.; Puschmann, A.; Mandel, T.; Xue, C.; Kaindl, G.; Hudson, E.; Shirley, D. A. *Phys. Rev. A* **1992**, *46*, 3935.
- [71] Chen, C. T.; Ma, Y.; Sette, F. *Phys. Rev. A* **1989**, *40*, 6737.
- [72] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1980**, *13*, 999.

- [73] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1979**, *12*, 137.
- [74] Francis, J. T.; Enkvist, C.; Lunell, S.; Hitchcock, A. P. *Can. J. Phys.* **1994**, *72*, 879–884.
- [75] Adachi, J.-i.; Kosugi, N.; Shigemasa, E.; Yagishita, A. *Chem. Phys. Lett.* **1999**, *309*, 427–433.
- [76] Hitchcock, A. P.; Brion, C. E. *J. Electron. Spectrosc. Relat. Phenom.* **1979**, *15*, 201–206.
- [77] Domke, M.; Xue, C.; Puschmann, A.; Mandel, T.; Hudson, E.; Shirley, D. A.; Kaindl, G. *Chem. Phys. Lett.* **1990**, *173*, 122–128.
- [78] Nayandin, O.; Kukk, E.; Wills, A.; Langer, B.; Bozek, J.; Canton-Rogan, S.; Wiedenhoeft, M.; Cubaynes, D.; Berrah, N. *Phys. Rev. A* **2001**, *63*, 062719.
- [79] Bodeur, S.; Milli, P.; Nenner, I. *Phys. Rev. A* **1990**, *41*, 252.
- [80] Gedat, E.; Püttner, R.; Domke, M.; Kaindl, G. *J. Chem. Phys.* **1998**, *109*, 4471.
- [81] Hudson, E.; Shirley, D. A.; Domke, M.; Remmers, G.; Kaindl, G. *Phys. Rev. A* **1994**, *49*, 161–175.
- [82] Cavell, R. G.; Jürgensen, A. *J. Electron. Spectrosc. Relat. Phenom.* **1999**, *101103*, 125–129.
- [83] Bodeur, S.; Esteva, J. M. *Chem. Phys.* **1985**, *100*, 415–427.
- [84] Nakata, A.; Imamura, Y.; Nakai, H. *J. Chem. Theory Comput.* **2007**, *3*, 1295–1305.
- [85] Saue, T. *ChemPhysChem* **2011**, *12*, 3077–3094.
- [86] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [87] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207–8215.
- [88] Song, J.-W.; Watson, M. A.; Nakata, A.; Hirao, K. *J. Chem. Phys.* **2008**, *129*, 184113.
- [89] Henderson, T. M.; Izmaylov, A. F.; Scuseria, G. E.; Savin, A. *J. Chem. Phys.* **2007**, *127*, 221103.
- [90] Henderson, T. M.; Izmaylov, A. F.; Scuseria, G. E.; Savin, A. *J. Chem. Theory Comput.* **2008**, *4*, 1254–1262.
- [91] Casida, M. E.; Gutierrez, F.; Guan, J.; Gadea, F.-X.; Salahub, D.; Daudey, J.-P. *J. Chem. Phys.* **2000**, *113*, 7062.
- [92] Di Mauro, E.; Hollenberg, C. P. *Adv. Mater.* **1993**, *5*, 384–386.
- [93] Niemeyer, C. M. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 585–587.
- [94] Niemeyer, C. M. *Angew. Chem. Int. Ed.* **2001**, *40*, 4128–4158.
- [95] Song, B.; Cuniberti, G.; Sanvito, S.; Fang, H. *Appl. Phys. Lett.* **2012**, *100*, 063101.
- [96] Langridge, R.; Marmur, J. *Science* **1964**, *143*, 1450–1451.
- [97] Sundaralingam, M. *Ann. N. Y. Acad. Sci.* **1975**, *255*, 3–42.
- [98] Camerman, N.; Camerman, A. *Science* **1968**, *160*, 1451–1452.
- [99] Davies, D. R. *Annu. Rev. Biochem.* **1967**, *36*, 321–364.

- [100] Kirtley, S. M.; Mullins, O. C.; Chen, J.; van Elp, J.; George, S. J.; Chen, C. T.; O'Halloran, T.; Cramer, S. P. *Biochim. Biophys. Acta.* **1992**, *1132*, 249–254.
- [101] Seifert, S.; Gavrilu, G. N.; Zahn, D. R. T.; Braun, W. *Surf. Sci.* **2007**, *601*, 2291–2296.
- [102] Yamada, T.; Shirasaka, K.; Takano, A.; Kawai, M. *Surf. Sci.* **2004**, *561*, 233–247.
- [103] Fujii, K.; Akamatsu, K.; Muramatsu, Y.; Yokoya, A. *Nucl. Instr. Meth. Phys. Res.* **2003**, *199*, 249–254.
- [104] Fujii, K.; Akamatsu, K.; Yokoya, A. *J. Phys. Chem. B* **2004**, *108*, 8031–8035.
- [105] Mochizuki, Y.; Koide, H.; Imamura, T.; Takemiya, H. *J. Synchrotron Radiat.* **2001**, *8*, 1003–1005.
- [106] MacNaughton, J.; Moewes, A.; Kurmaev, E. Z. *J. Phys. Chem. B.* **2005**, *109*, 7749–7757.
- [107] Ekström, U.; Norman, P.; Carravetta, V.; Ågren, H. *Phys. Rev. Lett.* **2006**, *97*, 143001.
- [108] Bolognesi, P.; Mattioli, G.; O'Keeffe, P.; Feyer, V.; Plekan, O.; Ovcharenko, Y.; Prince, K. C.; Coreno, M.; Amore Bonapasta, A.; Avaldi, L. *J. Phys. Chem. A.* **2009**, *113*, 13593–13600.
- [109] Healion, D. M.; Schweigert, I. V.; Mukamel, S. *J. Phys. Chem. A* **2008**, *112*, 11449–11461.
- [110] Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Trofimov, A. B.; Gromov, E. V.; Zaytseva, I. L.; Schirmer, J. *Chem. Phys.* **2008**, *347*, 360–375.
- [111] Wenzel, J.; Wormit, M.; Dreuw, A. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.
- [112] Robin, M. B. *Chem. Phys. Lett.* **1975**, *31*, 140–144.
- [113] Maganas, D.; Kristiansen, P.; Duda, L. C.; Gericke, A. K. *J. Phys. Chem. C* **2014**, *118*, 20163–20175.
- [114] Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.
- [115] Dyall, K. G. *J. Chem. Phys.* **1997**, *106*, 9618–9626.
- [116] Dyall, K. G. *J. Chem. Phys.* **2001**, *115*, 9136–9143.
- [117] Kutzelnigg, W.; Liu, W. *J. Chem. Phys.* **2005**, *123*, 241102.
- [118] Liu, W.; Peng, D. *J. Chem. Phys.* **2009**, *131*, 031104.
- [119] Zou, W.; Filatov, M.; Cremer, D. *Theor. Chim. Acta* **2011**, *130*, 633–644.
- [120] Cheng, L.; Gauss, J. *J. Chem. Phys.* **2011**, *135*, 084114.

Table 2.4: Calculated (B3LYP/def2-TZVP) and experimental thymine oxygen, nitrogen, and carbon core excitation energies (ω_{fi} , in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_h) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_p).

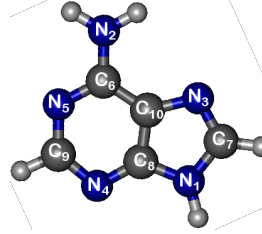


		OCDFT			Experiment	
ϕ_h	ϕ_p	ω_{fi}	f_{rel}	Peak	ω_{fi}	
Oxygen K-Edge						
O ₂	81.8% π_1^*	531.1	1.000	A	531.4	
O ₁	64.0% π_2^*	532.1	0.995	B	532.3	
O ₁	71.2% π_1^*	533.4	0.130	B'	533.8	
O ₂	78.3% π_2^*	533.8	0.143			
O ₁	69.5% D ₃	535.5	0.092	C	535.7	
O ₁	76.0% π_3^*	536.1	0.212			
O ₂	69.0% π_3^*	536.2	0.090			
O ₁	83.5% D ₂	537.1	0.040	D	537.1	
O ₁	35.6% D ₅	537.5	0.049			
O ₂	44.0% D ₅	537.7	0.065			
Nitrogen K-Edge						
N ₄	81.8% π_1^*	401.2	0.998	A	401.7	
N ₃	64.0% π_2^*	401.8	0.802			
N ₄	78.3% π_2^*	402.5	0.087	B	402.7	
N ₃	77.0% D ₁	403.1	0.895			
N ₄	65.9% D ₁	403.3	0.790			
N ₄	44.1% D ₃	404.2	0.894	C	404.1	
N ₃	69.5% D ₃	405.1	0.370	D	405.5	
N ₄	86.3% D ₂	405.3	0.852			
N ₃	71.2% π_1^*	405.7	0.502			
Carbon K-Edge						
C ₈	92.1% π_1^*	284.9	0.328	A	284.9	
C ₇	95.9% π_1^*	286.0	0.635	B	285.9	
C ₈	97.6% π_2^*	287.3	0.140	C	287.8	
C ₆	81.8% π_1^*	287.7	0.770			
C ₉	89.9% π_1^*	287.9	0.108			
C ₉	87.7% D ₁	288.5	0.205	D	288.4	
C ₅	64.0% π_2^*	289.1	1.000	E	289.4	
C ₉	63.1% π_3^*	289.3	0.294			
C ₉	49.5% D ₂	289.3	0.343			
C ₈	33.3% D ₄	290.3	0.098	F	290.7	
C ₉	30.5% D ₄	290.4	0.042			
C ₇	71.6% D ₃	290.4	0.041			

Table 2.5: Comparison of the most intense transitions at each edge of the thymine and adenine spectra. OCDFT assignments and energies (in eV) are shown and are compared to the most intense transitions from the theoretical ADC(2) spectra reported in Ref 110.

K-Edge	OCDFT	ADC(2)
Thymine		
Oxygen	$O_2 \rightarrow \pi_1^*$ (531.1)	$O_2 \rightarrow \pi_1^*$ (531.4)
Nitrogen	$N_4 \rightarrow \pi_1^*$ (401.2)	$N_4 \rightarrow \pi_1^*$ (401.5)
Carbon	$C_5 \rightarrow \pi_2^*$ (289.1)	$C_5 \rightarrow \pi_2^*$ (289.7)
Adenine		
Nitrogen	$N_5 \rightarrow \pi_2^*$ (399.4)	$N_5 \rightarrow \pi_2^*$ (399.6)
Carbon	$C_6 \rightarrow \pi_1^*$ (287.3)	$C_6 \rightarrow \pi_1^*$ (287.4)

Table 2.6: Calculated (B3LYP/def2-TZVP) and experimental adenine nitrogen and carbon core excitation energies (ω_{fi} , in eV) and relative oscillator strengths (f_{real}). For each calculated transition, we also report the label of the core atomic orbital (ϕ_{h}) and the largest ground-state virtual orbital contribution to the particle orbital (ϕ_{p}).



		OCDFT			Experiment	
ϕ_{h}	ϕ_{p}	ω_{fi}	f_{rel}	Peak	ω_{fi}	
Nitrogen K-Edge						
N ₄	81.0% π_1^*	399.1	0.825			
N ₃	63.7% π_1^*	399.3	0.915	A	399.5	
N ₅	92.6% π_2^*	399.4	1.000			
N ₅	92.4% π_1^*	399.7	0.001	A'	400.4	
N ₄	98.9% π_2^*	400.4	0.019			
N ₃	81.6% π_2^*	401.2	0.094	B'	401.3	
N ₂	82.1% π_1^*	401.4	0.351			
N ₁	69.3% π_1^*	401.8	0.587	B	401.9	
N ₂	78.4% D ₂	402.3	0.204			
N ₁	74.8% π_2^*	403.1	0.109			
N ₁	87.2% D ₁	403.3	0.418	C	403.0	
N ₂	57.1% D ₃	403.3	0.925			
Carbon K-Edge						
C ₁₀	92.6% π_2^*	286.3	0.270	A	286.4	
C ₉	81.0% π_1^*	286.5	0.924			
C ₁₀	92.4% π_1^*	286.7	0.242	B	286.8	
C ₇	82.1% π_1^*	286.9	0.878			
C ₆	69.3% π_1^*	287.3	1.000	C	287.4	
C ₈	63.7% π_1^*	287.4	0.965			
C ₁₀	34.7% π_3^*	287.9	0.083	C'	288.0	
C ₉	98.9% π_2^*	288.0	0.021			
C ₆	74.8% π_2^*	288.8	0.006			
C ₁₀	36.4% D ₂	289.2	0.033	D	289.0	
C ₈	66.2% π_3^*	289.2	0.012			
C ₉	77.4% π_3^*	289.4	0.010			
C ₇	87.1% π_2^*	289.4	0.029	E		
C ₇	57.1% D ₃	289.7	0.291			
C ₉	80.2% D ₁	290.0	0.225			
C ₈	90.2% D ₁	290.1	0.038	F		
C ₉	83.4% D ₂	290.1	0.140			
C ₇	94.8% π_3^*	290.4	0.073			
C ₆	77.2% π_3^*	290.7	0.052	G		
C ₇	56.6% D ₁	290.8	0.047			

Chapter 3

Predicting Near Edge X-ray Absorption Spectra with the Spin-Free Exact-Two-Component Hamiltonian and Orthogonality Constrained Density Functional Theory

“There is one topic I was NOT sorry to skip: the relativistic wave equation of Dirac”

Steven Weinberg

Chapter Abstract

Orthogonality constrained density functional theory (OCDFT) provides near edge X-ray absorption (NEXAS) spectra of first row elements within one electron volt from experimental values. However, with increasing atomic number, scalar relativistic effects become the dominant source of error in a nonrelativistic OCDFT treatment of core-valence excitations. In this work we report a novel implementation of the spin-free exact-two-component (X2C) one-electron treatment of scalar relativistic effects, and its combination with a recently developed OCDFT approach to compute a manifold of core-valence excited states. The inclusion of scalar relativistic effects in

OCDFT reduces the mean absolute error of second-row elements core-valence excitations from 10.3 to 2.3 eV. For all the excitations considered, the results from X2C calculations are also found to be in excellent agreement with those from low-order spin-free Douglas–Kroll–Hess relativistic Hamiltonians. The X2C-OCDFT NEXAS spectra of three organotitanium complexes (TiCl_4 , TiCpCl_3 , TiCp_2Cl_2) are in very good agreement with *unshifted* experimental results, and show a maximum absolute error of 5–6 eV. In addition, a decomposition of the total transition dipole moment into partial atomic contributions is proposed and applied to analyze the nature of the Ti pre-edge transitions in the three organotitanium complexes.

3.1 Introduction

The emergence of high-intensity and tunable synchrotron radiation sources has made X-ray absorption techniques some of the most powerful and versatile spectroscopic tools for elucidating local chemical information. Near-Edge X-ray Absorption Spectroscopy (NEXAS) probes electronic transitions from core to unoccupied states below the continuum,¹ providing a mean to investigate various electronic structure properties of the absorbing atom, such as oxidation state, coordination number, and spin/state symmetry.^{2–5} NEXAS has been applied to a wide range of chemical systems, including: small molecules in the gas phase⁶, molecules in liquid environments,^{7,8} large biological systems⁹, thin-films¹⁰, and semi-conducting materials¹¹. Experimental NEXAS studies are often coupled with computational simulations in order to quantify more subtle electronic structure information like discrete contributions to peak features, covalency, and orbital mixing.^{1,3,5,12,13}

Computing NEXAS spectra can be challenging since relativistic, relaxation, and correlation effects play a vital role in determining the accuracy of core-excitation energies. Several attempts have been made to extend many-body methods like coupled cluster theory^{14–20}, configuration interaction^{21–25}, and Green’s-function ap-

proaches^{26,27} to compute NEXAS spectra. NEXAS spectra can also be simulated using computationally less expensive methods such as multiple scattering theory^{28–30}, Slater transition state³¹, transition potential^{32,33}, static exchange method³⁴, configuration interaction singles^{35,36}, Δ SCF³⁷, linear response time-dependent density functional theory (TD-DFT)^{38–40}, and real-time TD-DFT^{41,42}.

Due to their low computational cost and ability to treat multiple excited states, methods based on TD-DFT are widely used to compute NEXAS. The accuracy of excitation energies computed with TD-DFT are strongly dependent on the choice of exchange-correlation (xc) kernel. Its failures are often attributed to errors in the Kohn–Sham eigenvalues, lack of frequency dependence in the xc-kernel, and the adiabatic approximation⁴³. For core excitations, TD-DFT provides only reliable relative energies⁴⁴ and empirical energy shifts must be used to achieve accurate absolute excitation energies.

Errors associated with the xc kernel and the Kohn–Sham orbital energies^{45,46} can be reduced by using time-independent DFT methods^{44,47,48}. We have recently proposed orthogonality constrained DFT (OCDFT),⁴⁹ a variational time-independent approach to compute excited states. OCDFT builds excited states via a generalized Kohn–Sham (KS) formalism that enforces additional orthogonality constraints between the ground and excited state wave functions. Our preliminary study⁵⁰ on a test set of 40 core-excited states showed that OCDFT produces absolute excitation energies that are more accurate than TDDFT. For example, using the popular B3LYP functional and a quadruple- ζ basis set, the mean absolute error for OCDFT and TDDFT is 1.0 and 21.6 eV, respectively. Our study found that the performance of nonrelativistic OCDFT was excellent for first-row elements. However, for second-row elements, relativistic effects were found to be the dominant source of error, and undoubtedly become more important when considering heavier elements. In this study we investigate the issue of relativistic effects in the prediction of K-edge spectra

and propose an inexpensive and accurate approach that combines scalar relativistic Hamiltonians with OCDFT.

Since a many-electron relativistic Hamiltonian cannot be easily derived from quantum electrodynamics, the theoretical treatment relativistic effects in electronic structure calculations is still an open problem.^{51–55} The simplest strategy is to decompose the relativistic effects into different contributions, i.e., scalar relativistic effects, spin-orbit coupling, magnetic induction, retardation effects, and finite nuclei, and include only those components relevant to the system and property under consideration^{56–60}. For K-edge X-ray absorption spectroscopy, scalar relativistic effects due to 1s orbital contraction are the dominant contribution, while other effects (e.g. spin-orbit coupling) become dominant when considering transitions from p and d orbitals.^{51,58,61,62}

A convenient approach to account for scalar relativistic effects consists in using the Foldy–Wouthuysen^{63,64} (FW) unitary transformation. The FW transformation can be used to reduce the four-component Dirac Hamiltonian to an exact or approximate quasi-relativistic 2-component (2C) Hamiltonian. The FW transformation may also be accompanied by elimination of the spin-dependent (vector) term from the 2C Hamiltonian to give a spin-free (SF) formalism. Several approaches that follow this philosophy exist: the zeroth order regular approximation (ZORA)^{65–67}, the Douglas–Kroll–Hess method^{68–72}, and various infinite-order schemes, including the normalized elimination of the small component,^{73–76} the Barysz–Sadlej–Snijde method⁷⁷, the infinite-order two-component scheme^{78,79}, and the exact-2-component approach (X2C).^{80–86} These methods differ in the way the FW transformation is performed: ZORA approximates it, the DKH method relies on a perturbative approach, and the infinite-order methods perform an exact transformation (within the computational basis). Since the ZORA Hamiltonian is not gauge invariant with respect to the energy scale,^{87–89} it does not yield accurate orbital energies and is therefore unsuitable for computing NEXAS spectra. The gauge dependency problem can be

fixed by using a scaled variant of ZORA^{88,90} or ZORA with a model potential⁸⁷. The former is designed to match the Dirac energies of hydrogen-like atoms, while the latter approach uses a model effective nuclear potential in the ZORA operator.

In this work we report a new implementation of the spin-free scalar quasi-relativistic X2C Hamiltonian at the one-electron level in the PSI4 *ab initio* quantum chemistry package⁹¹. A brief outline of the working equations of X2C is provided along with details of our implementation. We showcase the utility of X2C by characterizing its effect on the computation of OCDFT core excitation energies. The accuracy of NEXAS spectra computed using only scalar relativistic effects will be assessed by evaluating the performance of X2C-OCDFT over a test set of 37 unique core excitations from 13 different molecules spanning the first and second row of the periodic table. The performance of the X2C Hamiltonian will also be compared with that of second-, third-, and fourth-order DKH. The full treatment of relativistic effects enables us to study the NEXAS spectra of second- and third-row elements, which we demonstrate by computing the Ti K-edge of three organotitanium complexes.

3.2 Theory

3.2.1 One-Electron Spin-Free X2C

The central idea of the X2C method is to find a transformation that separates the positive (electronic) and negative solutions of the Dirac equation. More precisely, one seeks a unitary operator U that decouples the positive-energy (h_{++}^{FW}) and negative-energy (h_{--}^{FW}) blocks of the Dirac Hamiltonian (h^{D}):

$$U^\dagger h^{\text{D}} U = U^\dagger \begin{pmatrix} h_{LL} & h_{LS} \\ h_{SL} & h_{SS} \end{pmatrix} U = \begin{pmatrix} h_{++}^{\text{FW}} & 0 \\ 0 & h_{--}^{\text{FW}} \end{pmatrix} \quad (3.1)$$

The transformation U is built from the solutions of the Dirac equation. The one-electron treatment of X2C neglects the electron-electron interaction, and thus, it is only necessary to solve the one-electron Dirac equation to find h_{++}^{FW} . In a kinetically

balanced basis treatment^{92,93} in which the large component is expanded using an atomic orbital basis of dimension N , $\{\chi_\mu, \mu = 1, \dots, N\}$, the spin-free modified Dirac equation reads:

$$\begin{pmatrix} V & T \\ T & \frac{W^{\text{SF}}}{4c^2} - T \end{pmatrix} \begin{pmatrix} C^{\text{L}} \\ C^{\text{S}} \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & \frac{T}{2c^2} \end{pmatrix} \begin{pmatrix} C^{\text{L}} \\ C^{\text{S}} \end{pmatrix} E \quad (3.2)$$

where c is the speed of light, S is the overlap integral, and T , V , and W^{SF} are the matrix representation of the nonrelativistic kinetic energy ($\hat{p}^2/2$), the nuclear-electron attraction potential (\hat{V}), and the spin-free relativistic potential [$\hat{p} \cdot (\hat{V}\hat{p})$], respectively. The integrals $W_{\mu\nu}^{\text{SF}} = \langle \chi_\mu | \hat{p} \cdot (\hat{V}\hat{p}) | \chi_\nu \rangle$ can be easily computed as derivatives of the nuclear-electron attraction integrals with respect to nuclear coordinates. Note that the spin-free formulation of X2C simplifies the form of the relativistic potential and leads to a problem that has a lower dimensionality than the full Dirac equation: the matrices that enter the modified Dirac equation [Eq. (3.2)] have dimension $N \times N$, which leads to a $2N \times 2N$ eigenvalue problem.

In the X2C treatment,^{73,74,80–86} the positive-energy block of the Hamiltonian h_{++}^{FW} is given by the sum of a transformed kinetic (T_{X2C}) and potential energy (V_{X2C}) contribution, defined as

$$T_{\text{X2C}} = R^\dagger (TX + X^\dagger T - X^\dagger TX) R \quad (3.3)$$

and

$$V_{\text{X2C}} = R^\dagger (V + \frac{1}{4c^2} X^\dagger W^{\text{SF}} X) R \quad (3.4)$$

The coupling matrix $X = C^{\text{S}}(C^{\text{L}})^{-1}$ is obtained from the large (C^{L}) and small (C^{S}) components of the N positive energy solutions of the Dirac equation. The renormalization matrix⁸⁶

$$R = S^{-1/2} (S^{-1/2} \tilde{S} S^{-1/2})^{-1/2} S^{1/2}, \quad (3.5)$$

depends on the modified overlap matrix

$$\tilde{S} = S + \frac{1}{2c^2} X^\dagger TX. \quad (3.6)$$

Existing nonrelativistic electronic structure code can be extended to include scalar relativistic effects treated with the X2C method by replacing nonrelativistic kinetic and potential energy with the corresponding X2C operators T_{X2C} and V_{X2C} . The X2C one-electron operator is formed before the beginning of the HF/KS-SCF cycle and added to two-electron HF/KS operator. Our X2C implementation in PSI4 closely follows that reported by Cheng and Gauss⁹⁴ in their work on X2C analytic energy gradients.

3.2.2 Orthogonality Constrained DFT

Here we will give a brief introduction to orthogonality constrained density functional theory. Extensive details of the OCDFT formalism and its performance to calculate valence-valence excitation energies can be found in ref 49, while details on its extension to treat core excitations and the algorithm used to calculate multiple excited states can be found in ref 50. OCDFT is a variational time independent (TI) formulation of DFT that builds upon the approach developed by Ayers, Levy, and Nagy⁹⁵, where all n electronic states $\Psi^{(n)}$ of an N -electron system have a unique corresponding density functional $E^{(n)}[\rho]$, which is a generalization of the ground state functional $E^{(0)}[\rho]$. The energy functional for state n minimizes the expectation values of the energy while imposing that the trial wave function (Ψ) be compatible with the density (ρ) and orthogonal to the first n states:

$$E^{(n)}[\rho] = \min_{\substack{\Psi \rightarrow \rho \\ \Psi \perp \{\Psi^{(k)}\}}} \langle \Psi | \hat{H} | \Psi \rangle, k = 0, 1, \dots, n-1 \quad (3.7)$$

OCDFT provides a practical realization of the Ayers–Levy–Nagy TIDFT approach that is based on a generalized Kohn–Sham scheme. In OCDFT, to each electronic state $\Psi^{(n)}$ corresponds an auxiliary system Ψ of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho^{(n)}$. The wave function for the auxiliary system is a single Slater determinant, $|\Phi^{(n)}\rangle = |\phi_1^{(n)} \phi_2^{(n)} \cdots \phi_N^{(n)}\rangle$, where the set of orbitals $\{\phi_i^{(n)}\}$ are different for each electronic state. In the derivation of a practical Kohn–Sham scheme

for OCDFT, the energy functional is augmented with an orthogonality condition on the auxiliary wave functions

$$\langle \Phi^{(m)} | \Phi^{(n)} \rangle = \delta_{mn} \quad \forall m, n \quad (3.8)$$

This orthogonality condition prevents the excited state from optimizing down to the ground state solution, effectively avoiding the variational collapse problems that plague most Δ SCF methods. In the resulting variational KS method, every excited state has an associated energy functional $E_{\text{KS}}^{(n)}[\{\phi_i^{(n)}\}]$ which has a unique exchange-correlation (xc) contribution $E_{\text{xc}}^{(n)}$.

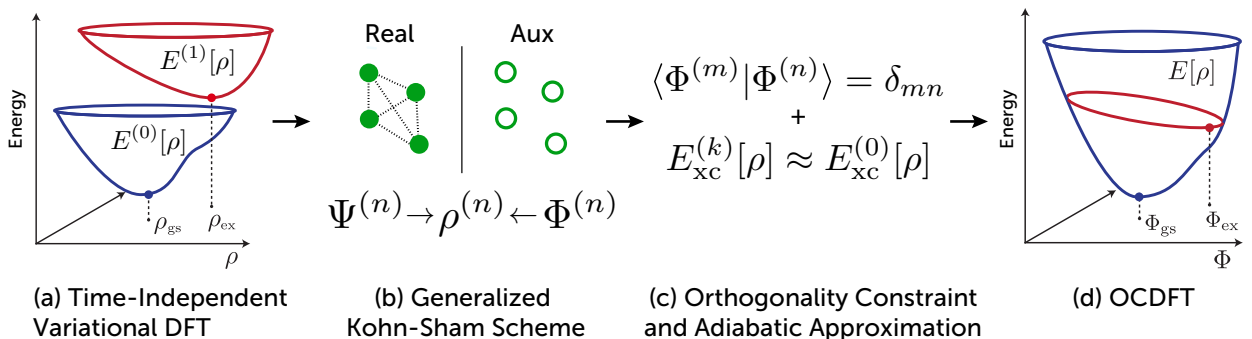


Figure 3.1: Outline of Orthogonality Constrained DFT: (a) In the Ayers–Levy–Nagy time-independent variational formulation of DFT each electronic is associated with a unique density functional. (b) For each electronic state $\Psi^{(n)}$, an auxiliary system of non-interacting electrons with wave function $\Phi^{(n)}$ is introduced. $\Psi^{(n)}$ and $\Phi^{(n)}$ are assumed to have the same electron density $\rho^{(n)}$. (c) We then impose explicit orthogonality conditions onto the auxiliary wave functions and approximate the unknown exchange-correlation (xc) functional $E_{\text{xc}}^{(n)}$ with the ground-state functional. (d) The final OCDFT functional consists in the ground state functional augmented with an orthogonality constraint.

As illustrated in Figure 3.1, OCDFT invokes an adiabatic approximation similar to the one adopted in TDDFT, in which the xc contribution for each excited state is approximated by the ground state functional $E_{\text{xc}}^{(0)}$. This results in the OCDFT

functional for excited state n , which is given by:

$$\begin{aligned}
 E_{\text{OCDFT}}^{(n)}[\{\phi_i^{(n)}\}] = & -\frac{1}{2} \sum_i^{\text{occ}} \langle \phi_i^{(n)} | \nabla^2 | \phi_i^{(n)} \rangle \\
 & + \int d\mathbf{r} v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) \\
 & + E_{\text{Coul}}[\rho^{(n)}] + E_{\text{xc}}^{(0)}[\rho^{(n)}]
 \end{aligned}
 \tag{3.9}$$

where $v_{\text{ext}}(\mathbf{r})$ is the external potential and $E_{\text{Coul}}[\rho^{(n)}]$ is the classical electron-electron interaction energy term.

In this work we solve the OCDFT equations via our constrained multiple hole/particle (CMHP) algorithm.⁵⁰ Within the CMHP scheme, each excited state is characterized by a pair of hole (ϕ_h^n) and particle (ϕ_p^n) orbitals, which are required to span the occupied and virtual space of the ground state wave function, respectively. These may be interpreted as the orbitals from which an electron is annihilated and created as a result of an electronic excitation. Rather than solving for the minimum orthogonality conditions between the states, the CMHP algorithm exploits the locality of hole orbitals in core-excited states by imposing sufficient *orbital orthogonality conditions*. This approximation simplifies the Lagrangian, producing a computationally feasible minimization procedure. In addition, it allows us to readily compute multiple excited states by enforcing mutual orthogonality conditions between the hole and particle orbitals of each subsequent excited state. At the same time, the CMHP scheme fully accounts for relaxation of the remaining spectator orbitals.

3.2.3 Comparison of OCDFT and TDDFT for Core Excitations

We have previously shown that OCDFT yields core excitation energies that are in better agreement with experiment than TDDFT.⁵⁰ Specifically, we find that OCDFT excitation energies either do not require shifting, or if an energy shift is necessary, it is significantly smaller than the typical values used in TDDFT. In this section we address the following questions: 1) Why do TDDFT core-excitation spectra require

larger energy shifts than OCDFT? and, 2) How much do *shifted* TDDFT and OCDFT core-excitation spectra differ?

To showcase the differences between TDDFT and OCDFT, we computed the carbon and oxygen K-edge spectra of CO using a series of functionals with increasing amount of Hartree–Fock exchange. For comparison, we also include results from time-dependent Hartree–Fock (TDHF) and orthogonality-constrained Hartree–Fock (OCHF). In this example, we will not consider the effect of relativity since for C and O these contribute to shifts in the excitation energies of the order of 0.1–0.3 eV (see Section 3.4). We will denote the fraction of Hartree–Fock exchange included in a functional with the letter δ , and for simplicity we exclude range-corrected functionals from our analysis.

Figure 3.2 shows the relative peak positions of the OCDFT and TDDFT spectra and the corresponding energy shift necessary to match the $C_{1s \rightarrow \pi^*}$ and $O_{1s \rightarrow \pi^*}$ peaks with their experimental values (287.40 and 534.10 eV for C and O, respectively). See the supplementary material for a table with the data used to create this figure. Note that for TDDFT, the C spectrum energy shift ranges from -6.96 (TDHF, $\delta = 1$) to $+16.24$ eV (BLYP, $\delta = 0$), while the O spectrum energy shift spans an even wider range (from -15.95 to $+21.97$ eV). In comparison, OCDFT shifts are smaller, and range from $+0.25$ eV (OCHF) to $+0.99$ eV (BLYP) for carbon, and from $+0.87$ eV (BHLYP) to $+1.77$ eV (OCHF) for oxygen.

Figure 3.2 also shows the relative position of the second, third, and fourth peaks of the K-edge spectrum. In TDDFT, the position of these three transitions is heavily dependent on the amount of Hartree–Fock exchange. In accordance with previous studies,⁹⁶ we note that only a functional with a large fraction of Hartree–Fock exchange like BHLYP ($\delta = 0.5$) can yield TDDFT relative excitation energies that are in sufficient agreement with experiment. This does not seem to be the case for OCDFT. Indeed, OCDFT relative peak positions are mostly insensitive to the choice

of the functional being used, and surprisingly, both functionals with no (BLYP) and 100% (HF) Hartree–Fock exchange yield C and O K-edge spectra that are in good agreement with the experimental spectra. Overall, this example shows that there are fundamental differences between core excitation energies predicted by TDDFT and OCDFT. In addition, there will also be differences in the oscillator strengths predicted by TDDFT and OCDFT.

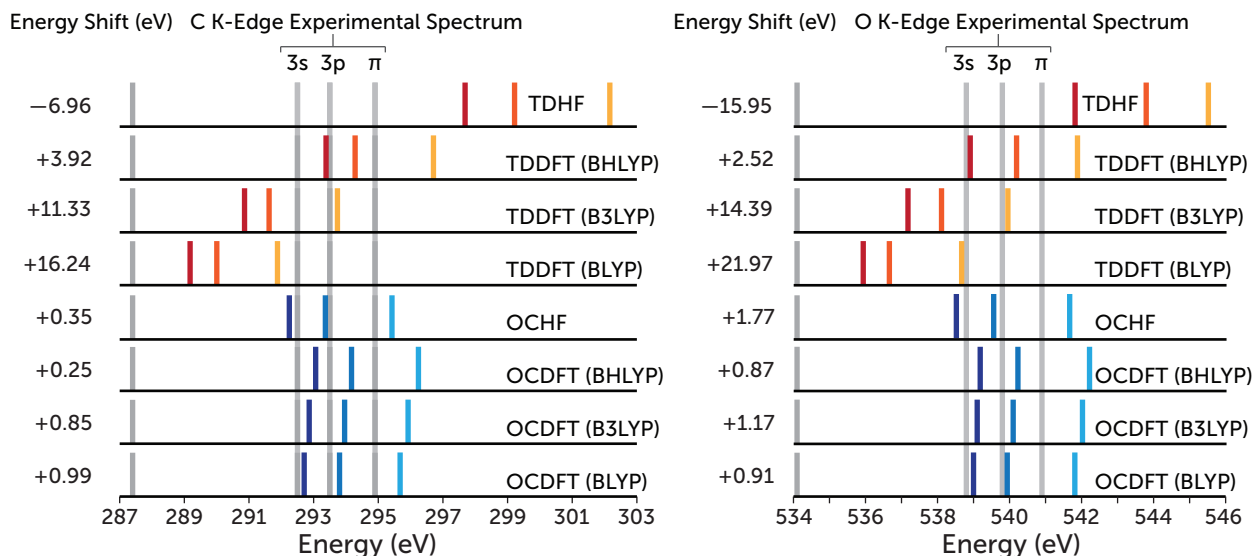


Figure 3.2: Relative Peak Positions of the C (left) and O (right) K-edge spectra of CO. Spectra are shifted by the amount required to align the first peak to the experimental peak values (287.40 and 534.10 eV for the C and O spectrum respectively). All calculations use the aug-cc-pVQZ basis set.

To understand the differences between TDDFT and OCDFT we analyze a model that captures the essential features of core-excited states. The model used here is similar to that considered by Casida et al.⁹⁷ and by Ziegler et al.⁹⁸. We consider a system with three orbitals (ϕ_i , ϕ_a , ϕ_b) of different spatial symmetry. The ground state of the model is assumed to be the closed-shell determinant $|(\phi_i)^2\rangle$, so that ϕ_a and ϕ_b are unoccupied orbitals. We model K-edge transitions as “charge-transfer-like” states, that is we assume that the core orbital ϕ_i has negligible spatial overlap with virtual orbitals. This assumption translates in to the condition $\int d\mathbf{r} |\phi_i(\mathbf{r})|^2 |\phi_p(\mathbf{r})|^2 = 0$ for ϕ_p equal to ϕ_a or ϕ_b .

Table 3.1: Singlet Excitation Energies for a Toy Model of Two Electrons in Three Molecular Orbitals.

Method	Singlet excitation energy ^a
CIS	$\omega_{ai}^{\text{CIS}} = h_a - h_i + J_{ai} - J_{ii}$ $\Delta\omega^{\text{CIS}} = \omega_{bi} - \omega_{ai} = h_b - h_a + J_{bi} - J_{ai}$
TDDFT	$\omega_{ai}^{\text{TDDFT}} = \omega_{ai}^{\text{CIS}} + (1 - \delta)(v_a^x - v_i^x + J_{ai} - J_{ii})$ $\Delta\omega^{\text{TDDFT}} = \Delta\omega^{\text{CIS}} + (1 - \delta)(v_b^x - v_a^x + J_{bi} - J_{ai})$
OCDFT	$\omega_{ai}^{\text{OCDFT}} = \omega_{ai}^{\text{CIS}} + (1 - \delta)\left(v_a^x - v_i^x + \frac{1}{2}J_{aa} - \frac{1}{2}J_{ii} + \frac{1}{2}f_{aa}^x + \frac{1}{2}f_{ii}^x\right)$ $\Delta\omega^{\text{OCDFT}} = \Delta\omega^{\text{CIS}} + (1 - \delta)\left(v_b^x - v_a^x + \frac{1}{2}J_{bb} - \frac{1}{2}J_{aa} + \frac{1}{2}f_{aa}^x - \frac{1}{2}f_{bb}^x\right)$

^aThe ground state wave function is given by $|\Phi\rangle = |(\phi_i)^2\rangle$. All energies are expressed in terms of one-particle energies, $h_p = \langle\phi_p|\hat{h}|\phi_p\rangle$, density functional exchange potential $v_p^x = \langle\phi_p|v_x|\phi_p\rangle$, Coulomb $J_{pq} = (pp|r_{12}^{-1}|qq)$, and exchange kernel $f_{pp}^x = (pp|f_x^{\alpha\alpha}|pp)$ integrals. The excitation energy of OCDFT is computed by a second-order expansion of the exchange-correlation functional around the ground state density. All contributions arising from the correlation functional were neglected and minimum overlap between involved orbital is assumed ($\langle|\phi_a|^2||\phi_i|^2\rangle \sim 0$). The amount of Hartree–Fock exchange included in the density functional is indicated with $\delta \in [0, 1]$.

We consider the singlet core excitation energies for the transitions $\phi_i \rightarrow \phi_a$ (ω_{ai}) and $\phi_i \rightarrow \phi_b$ (ω_{bi}), together with the relative excitation energy $\Delta\omega = \omega_{bi} - \omega_{ai}$. For this model, these excitation energies can be written down explicitly in terms of molecular integrals.^{97,98} The reference excitation energies for our model are computed using CIS and are reported in Table 3.1. In addition, we report the same quantities as obtained from TDDFT using the Tamm–Dancoff approximation and OCDFT with the exchange-correlation functional approximated as a Taylor expansion around the ground state density. We refer the reader to Table 3.1 for the definition of the integrals used in this section.

The origin of the large energy shift required to match TDDFT K-edge spectra with experiment may be traced down to the difference between the TDDFT and CIS

excitation energy:

$$\omega_{ai}^{\text{TDDFT}} - \omega_{ai}^{\text{CIS}} = (1 - \delta) (v_a^x - v_i^x + J_{ai} - J_{ii}). \quad (3.10)$$

In the local density approximation (LDA), the exchange potential is equal to $v^x[\rho] = -\frac{4}{3}C_x\rho^{1/3}$, where C_x is a positive constant. Therefore, we estimate that the matrix element $v_p^x = \langle \phi_p | v_x | \phi_p \rangle$ will be larger for orbitals localized near regions with high electron density (i.e. atomic nuclei). As a consequence, for core excitations the term $v_a^x - v_i^x$ will be positive. Indeed, for the lowest K-edge transition of CO we estimate this quantity from the full exchange-correlation potential as $v_a^{\text{xc}} - v_i^{\text{xc}}$, and find that it is about +48.4 eV for C and +70.9 eV for O.

Let us now turn to the contribution from the second term in Eq (3.10), $J_{ai} - J_{ii}$. This term can be interpreted physically as the sum of the energy released from breaking the core electron pair ($-J_{ii}$) and the interaction of the unpaired electrons in the hole and particle orbitals (J_{ai}). This term is already present with the correct prefactor in the CIS expression for the excitation energy (ω_{ai}^{CIS}), therefore, in the case of TDDFT it is double counted. For core excitations one has that $J_{ii} \gg J_{ai}$. Therefore, the second term in Eq (3.10) is negative, and has the net effect of shifting down the excitation energies. Since TDDFT core excitation energies are always underestimated, we expect that $J_{ai} - J_{ii}$ is larger in magnitude than the term $v_a^x - v_i^x$. Indeed, in the case of the lowest energy transition, we find that $J_{ai} - J_{ii}$ is respectively -89.7 and -123.4 eV for the C and O K-edge. Combining these estimates with those for the exchange contributions, we predict TDDFT energy shifts of the order of -41.4 eV for C and -52.1 eV for O. These estimates do not agree well with the TDDFT actual shifts, but they still capture some of the essential features of our data for CO: 1) they predict negative shifts that are dominated by the two-electron repulsion contribution, and 2) they predict that the oxygen energy shift is larger in magnitude than that required for carbon. Note that this model predicts that as the percentage of Hartree–Fock exchange is increased ($\delta \rightarrow 1$), the TDDFT core excitation energies

require less shifting. This observation is in agreement with numerical results. Indeed, one way to improve the accuracy of TDDFT for core excitation energies is to use functionals with a relatively high fraction of Hartree–Fock exchange (50–60%).⁹⁶

In the case of OCDFT, the equation for the energy shift differs from that of TDDFT and it is somewhat more difficult to analyze. In addition to a term proportional to $v_a^x - v_i^x$, which is also common to the TDDFT expression, the OCDFT shift has various contributions from several local self-interaction terms. Focusing on the contribution from the two-electron interactions, $\frac{1}{2}(J_{aa} - J_{ii})$, this quantity is equal to -61.9 and -45.0 eV for O and C, respectively. In comparison to the TDDFT two-electron contribution, the magnitude of this term is smaller and there is better cancellation of errors.

Turning to our second question, we use the three-state model to compute the relative excitation energies for both the TDDFT and OCDFT spectra. The excitation energy difference between the two single excitations as predicted by TDDFT ($\Delta\omega^{\text{TDDFT}} = \omega_{bi}^{\text{TDDFT}} - \omega_{ai}^{\text{TDDFT}}$) differs from the CIS value by the quantity $(1 - \delta)(v_b^x - v_a^x + J_{bi} - J_{ai})$. The most worrisome contribution to the error is the difference $J_{bi} - J_{ai}$, which effectively introduces a distance dependence in the excitation energy error.

In the case of OCDFT, the relative excitation energy differs from the CIS value by a combination of exchange potential integrals ($v_b^x - v_a^x$, which also contributes to the TDDFT error) and a sum of differences of self-interaction energies for orbitals ϕ_a and ϕ_b . Thus, all the terms that contribute to the OCDFT relative energy error are local, and relative excitation energies should not show a marked dependence on the distance between the hole and particle orbitals.

3.3 Computational Details

Core excitation energies are computed with OCDFT using the nonrelativistic Hamiltonian, the spin-free X2C Hamiltonian, and the second-, third-, and fourth-order variants of the DKH (DKH2–4) Hamiltonian. Benchmark core excitation energies were computed for 13 small molecules (CH_4 , C_2H_4 , HCN , CO , H_2CO , N_2 , N_2O , SO_2 , SiH_4 , PH_3 , H_2S , HCl , Cl_2). The geometries of the test set molecules were optimized with nonrelativistic DFT using Becke’s B3LYP^{99–102} exchange-correlation functional and the def2-QZVP^{103,104} basis set. Individual excitation energies are compared with data obtained from gas-phase NEXAS experiments^{105–120}.

Geometries for the three titanium complexes used in the study are taken from ref. 121 and 122. The first is a simple tetrahedral TiCl_4 complex, while the other two are the mono- and bis-cyclopentadienyl (Cp) complexes TiCpCl_3 and TiCp_2Cl_2 . The geometries used in our computations show only small deviations from crystal structures (changes of less than 0.02 Å in bond length and 0.5° in bond angle). OCDFT results for titanium complexes are obtained with the fully uncontracted cc-pVDZ^{123,124} (un-cc-pVDZ) basis set with the two-electron integrals approximated using a density fitting approximation^{125–129} with the def2-TZVP-JK auxiliary basis¹³⁰. For the test set, the fully uncontracted aug-cc-pVDZ¹³¹ (un-aug-cc-pVDZ) basis set was used without the density fitting approximation. Fully uncontracted basis sets are required to provide a proper description of the small component of the wave function and to guarantee a fair comparison between nonrelativistic and relativistic methods. For these organotitanium complexes, OCDFT core excitation energies are computed using only the X2C Hamiltonian.

X2C-OCDFT calculations are performed with the PSI4 *ab initio* quantum chemistry package⁹¹ using our implementation of X2C in PSI4. DKH+OCDFT calculations are performed with PSI4 using the code of Reiher, Wolf, and Hess^{68,71}. All TD-DFT

Table 3.2: Excitation energies of the three lowest energy 1s core excitations for each atom in CO calculated with X2C-OCDFT using the B3LYP functional and fully uncontracted cc-pVXZ and aug-cc-pVXZ basis sets ($X = D, T, Q$). Experimental values are taken from Refs. 107 and 110. Errors greater than 1 eV are indicated with bold font.

Transition	Exp.	un-cc-pVXZ			un-aug-cc-pVXZ		
		D	T	Q	D	T	Q
O _{1s} → π^*	534.1	-0.4	-0.8	-0.9	-0.4	-0.8	-0.9
O _{1s} →3s	538.8	3.9	1.9	0.9	0.0	-0.5	-0.6
O _{1s} →3p	539.8	7.8	4.1	2.4	0.1	-0.5	-0.6
C _{1s} → π^*	287.4	-0.5	-0.8	-0.8	-0.5	-0.8	-0.9
C _{1s} →3s	292.5	4.0	2.1	1.1	0.0	-0.4	-0.5
C _{1s} →3p	293.4	7.1	4.0	2.4	0.3	-0.2	-0.3
MAE		4.0	2.3	1.4	0.2	0.5	0.6

calculatons are done with the ORCA software package version 3.0.3¹³². Natural population analysis on the Ti complexes were performed using the JANPA¹³³ software package.

3.4 Results and Discussion

3.4.1 Calibration of X2C-OCDFT Core Excitation Energies

In order to gauge the importance of relativistic effects in OCDFT core excitation energies, nonrelativistic (NR) and relativistic computations are performed on a test set of 36 transitions from 13 molecules containing first- and second-row elements.

Table 3.2 shows the systematic convergence of X2C-OCDFT on a test case of six core excitations of CO using the fully uncontracted correlation consistent cc-pVXZ and aug-cc-pVXZ series of basis sets ranging from double- to quadruple- ζ quality. Both 1s $\rightarrow \pi^*$ excitations in CO are represented well across all the basis sets considered, with the addition of diffuse functions having no significant impact on the accuracy of the excitation energies compared to experiment. However, the cc-pVXZ series of basis sets struggle to properly describe the 1s \rightarrow 3s and 1s \rightarrow 3p transitions in CO, and the inclusion of diffuse functions drastically improves the accuracy of these

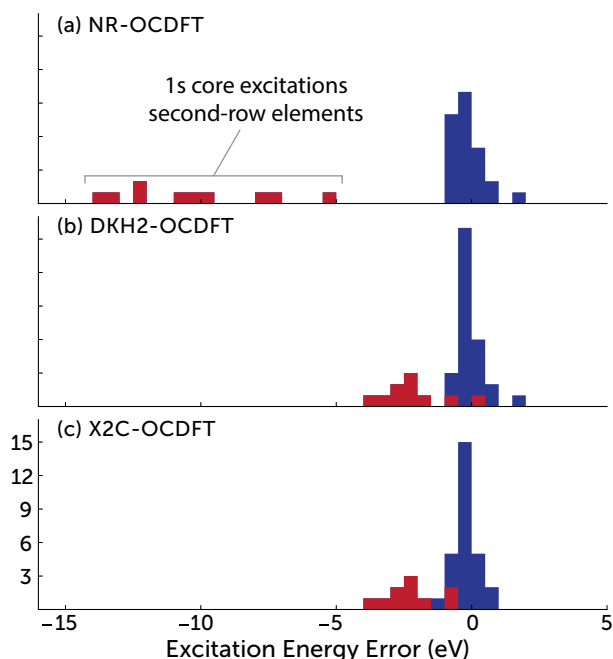


Figure 3.3: Error distribution in the computed core excited states for: (a) nonrelativistic OCDFT, (b) OCDFT with the second-order Douglas–Kroll–Hess Hamiltonian, and (c) OCDFT with the exact-2-component Hamiltonian. All calculations use the B3LYP functional and a fully uncontracted aug-cc-pVDZ basis set. Red bars indicate core excitations from 1s orbitals of second-row elements.

excitation energies. The un-aug-cc-pVDZ basis set provides a convenient compromise between accuracy and cost, and was chosen to benchmark the full test set of core excitations.

Table 3.3 and 3.4 report core excitations from first- and second-row elements. The performance of OCDFT for *first-row* core excitations is slightly improved upon inclusion of scalar relativistic effects. Going from NR-OCDFT to OCDFT with scalar relativistic Hamiltonians (X2C, DKH2–4) reduces the error by about 0.2 eV. In addition, the maximum absolute deviation between X2C-OCDFT and NR-OCDFT for first-row core excitations is only 0.4 eV ($O_{1s \rightarrow 3s}$ transition of N_2O). X2C and low-order DKH methods yield results that are identical, with negligible differences in MAE. This similarity in performance between NR-OCDFT and multiple scalar relativistic Hamiltonians is consistent with the general observation that scalar relativistic effects are not significant for the core excitation energies of first row elements.^{16,24,50} A

Table 3.3: OCDFT core excitation energies for molecules containing first-row elements using nonrelativistic (NR) and relativistic (X2C,DKH2, DKH3, DKH4) Hamiltonians. Computations were performed using the uncontracted aug-cc-pVDZ basis with the B3LYP exchange-correlation functional (E_{xc}^{B3LYP}) and 100% non-local exact exchange (E_x^{HF}). OCDFT results are reported here as deviations from the experimental excitation energies (in eV). The mean absolute error (MAE) is also reported for each method. Experimental values are taken from refs 105–114.

Molecule	Excitation	Exp.	Error					
			E_{xc}^{B3LYP}					E_x^{HF}
			NR	X2C	DKH2	DKH3	DKH4	X2C
CO	$C_{1s \rightarrow \pi^*}$	287.4	-0.6	-0.6	-0.5	-0.5	-0.5	0.1
	$C_{1s \rightarrow 3s}$	292.4	0.0	0.0	0.1	0.1	0.1	0.0
	$O_{1s \rightarrow \pi^*}$	534.2	-0.9	-0.5	-0.5	-0.5	-0.5	-1.1
	$O_{1s \rightarrow 3s}$	538.9	-0.5	-0.5	-0.1	-0.1	-0.1	-1.2
H_2CO	$O_{1s \rightarrow \pi^*}$	530.8	-0.6	-0.2	-0.2	-0.2	-0.2	-0.8
	$O_{1s \rightarrow 3s}$	535.4	-0.5	-0.2	-0.2	-0.2	-0.2	-1.3
	$C_{1s \rightarrow 3s}$	290.2	-0.3	-0.2	-0.2	-0.2	-0.2	-0.6
	$C_{1s \rightarrow \pi^*}$	286.0	-0.4	-0.3	-0.3	-0.3	-0.3	-0.2
N_2O	$O_{1s \rightarrow \pi^*}$	534.8	-0.5	-0.2	-0.1	-0.1	-0.1	-1.3
	$O_{1s \rightarrow 3s}$	536.7	-0.6	-0.2	-0.2	-0.2	-0.2	-3.1
	${}^cN_{1s \rightarrow 3s}^a$	407.5	0.2	0.3	0.3	0.3	0.3	39.2 ^b
	${}^tN_{1s \rightarrow \pi^*}^a$	401.1	-0.8	-0.5	-0.6	-0.6	-0.6	10.4 ^b
N_2	$N_{1s \rightarrow \pi^*}$	401.0	-0.8	-0.5	-0.5	-0.5	-0.5	8.9 ^b
HCN	$C_{1s \rightarrow \pi^*}$	286.4	-0.3	-0.2	-0.2	-0.2	-0.2	-0.3
	$C_{1s \rightarrow 3s}$	289.1	-0.2	-0.1	-0.1	-0.1	-0.1	-0.4
	$N_{1s \rightarrow \pi^*}$	399.7	-0.6	-0.4	-0.4	-0.4	-0.4	-0.6
	$N_{1s \rightarrow 3s}$	401.8	0.2	0.4	0.4	0.4	0.4	-1.0
CH_4	$C_{1s \rightarrow 3p}$	288.0	-0.3	-0.2	-0.2	-0.2	-0.2	-0.3
	$C_{1s \rightarrow 3s}$	287.1	-0.7	-0.6	-0.6	-0.6	-0.6	-0.6
C_2H_2	$C_{1s \rightarrow \pi^*}$	285.8	-0.5	-0.3	-0.4	-0.4	-0.4	7.6 ^b
MAE			0.5	0.3	0.3	0.3	0.3	0.8

^aThe subscripts c and t stand for the center and tail nitrogen of N_2O , respectively.

^bExcitation excluded from the computation of the MAE.

similar conclusion can be drawn from the second row 2p excitation energies shown in the bottom half of Table 3.4 where the difference in MAE between NR and relativistic OCDFT is negligible, with identical results produced across all orders of DKH.

Figure 3.3 shows the distribution of errors for all the test set transitions computed with the NR, X2C, and DKH2 Hamiltonians. Note how the 1s orbital (K-edge) tran-

Table 3.4: OCDFT core excitation energies for molecules containing first-row elements using nonrelativistic (NR) and relativistic (X2C,DKH2, DKH3, DKH4) Hamiltonians. Computations were performed using the uncontracted aug-cc-pVDZ basis with the B3LYP exchange-correlation functional (E_{xc}^{B3LYP}) and 100% non-local exact exchange (E_x^{HF}). OCDFT results are reported here as deviations from the experimental excitation energies (in eV). The mean absolute error (MAE) is also reported for each method. Experimental values are from refs [115–120](#)

Molecule	Excitation	Exp.	Error					
			E_{xc}^{B3LYP}					E_x^{HF}
			NR	X2C	DKH2	DKH3	DKH4	X2C
SiH ₄	Si _{1s} → σ^*	1842.5	-5.3	-1.0	-1.0	-1.0	-1.0	1.6
PH ₃	P _{1s} → σ^*	2145.8	-7.6	-1.8	-1.8	-1.8	-1.8	1.2
H ₂ S	S _{1s} → σ^*	2473.1	-10.5	-2.8	-2.8	-2.8	-2.8	0.7
	S _{1s} →4p	2476.3	-9.9	-2.2	-2.2	-2.2	-2.2	0.8
SO ₂	S _{1s} → π^*	2473.8	-10.6	-2.9	-2.9	-2.9	-2.9	0.5
	S _{1s} →4p	2478.4	-7.3	-0.8	0.4	0.4	0.4	3.7
HCl	Cl _{1s} → σ^*	2823.9	-13.2	-3.2	-3.2	-3.2	-3.2	0.4
	Cl _{1s} →4p	2827.8	-12.2	-2.2	-2.2	-2.2	-2.2	1.1
Cl ₂	Cl _{1s} → σ^*	2821.3	-13.6	-3.6	-3.6	-3.6	-3.6	16.8 ^a
	Cl _{1s} →4p	2828.5	-12.4	-2.3	-2.4	-2.3	-2.3	18.0 ^a
MAE			10.3	2.3	2.3	2.2	2.2	1.3
SiH ₄	Si _{2p} → σ^*	102.8	0.9	0.9	0.9	0.9	0.9	1.2
PH ₃	P _{2p} → σ^*	132.3	0.3	0.3	0.3	0.3	0.3	0.6
H ₂ S	S _{2p} → σ^*	164.5	0.7	0.7	0.7	0.7	0.7	1.8
H ₂ S	S _{2p} →4s	166.5	1.6	-1.5	1.6	1.6	1.6	1.8
Cl ₂	Cl _{2p} → σ^*	198.7	-0.5	-0.5	-0.5	0.0	0.0	-0.1
HCl	Cl _{2p} →4s	201.0	0.3	0.3	0.3	0.6	0.3	1.3
MAE			0.7	0.7	0.7	0.7	0.6	1.1

^aExcitation excluded from the computation of the MAE.

sitions for second-row elements (shaded red in Figure 3.3) are consistently underestimated by NR-OCDFT and yield a MAE of 10.3 eV. These troublesome transitions are drastically improved with the inclusion of scalar relativistic effects, reducing the MAE down to 2.3 eV. Even for excitations that involve second-row element we notice that the X2C and DKH2–4 methods are in excellent agreement (difference in energy less than 0.1 eV).

In order to test the importance of correlation effects for the computation of core

excitation energies, OCDFT was applied to the test set using a functional with 100% non-local exchange (E_x^{HF}). These results are displayed in the far right-hand column of Tables 3.3 and 3.4. It is important to notice that for transitions in molecules with equivalent or near-equivalent nuclei (C_2H_2 , N_2O , N_2 , Cl_2), we previously determined that it is necessary to seek a wave function with hole orbitals localized on one atom.⁵⁰ Broken symmetry calculations were used in order to obtain OCDFT excitation energies of C_2H_2 , N_2 , N_2O and Cl_2 . However, we were unable to obtain localized hole orbitals for these molecules with OCDFT and the E_x^{HF} functional. When transitions with delocalized hole orbitals are excluded, the HF functional provides results that are in modest agreement with those obtained using the B3LYP functional. For the first row elements the maximum difference observed in Table 3.3 between E_x^{HF} and $E_{\text{xc}}^{\text{B3LYP}}$ is 2.9 eV (N_2O , $\text{O}_{1s \rightarrow 3s}$ transition), while for second row elements the corresponding largest difference observed in Table 3.4 is 4.5 eV (SO_2 , $\text{S}_{1s \rightarrow 4p}$ transition).

Table 3.5 reports estimates of relativistic and correlation effects. The former are computed as the difference between the NR- and X2C-OCDFT excitation energies, while the latter are estimated as the difference between the relativistic results using the B3LYP and Hartree–Fock exchange. The results in Table 3.5 show that relativistic and correlation effects follow the trend: 1s first-row ; 2p second-row ; 1s second row. For first row 1s and second row 2p excitations, correlation effects are larger than relativistic effects; this trend reverses in the second row 1s excitations.

3.4.2 Ti K-Edge NEXAS of Organotitanium Complexes

In this section we will apply the X2C-OCDFT approach to compute the Ti K-edge spectrum of TiCl_4 and its mono- and bis-cyclopentadienyl derivatives, TiCpCl_3 and TiCp_2Cl_2 . These compounds have drawn ample attention due to their catalytic activity and potential use as anticancer drugs. The increased contribution of relativistic effects in the Ti core and the closed shell electronic configuration (d^0) of tetracoor-

Table 3.5: Importance of relativistic and correlation effects in core excitations computed with X2C-OCDFT. Relativistic effects are estimated as the difference between NR and X2C OCDFT calculations [$\text{OCDFT}^{\text{NR}}(E_{\text{xc}}^{\text{B3LYP}}) - \text{OCDFT}^{\text{X2C}}(E_{\text{xc}}^{\text{B3LYP}})$], while correlation effects are estimated as the difference between B3LYP and HF OCDFT calculations [$\text{OCDFT}^{\text{X2C}}(E_{\text{xc}}^{\text{B3LYP}}) - \text{OCDFT}^{\text{X2C}}(E_{\text{x}}^{\text{HF}})$]. Molecules with symmetry equivalent nuclei were excluded from this analysis due to inability of the HF functional to obtain a broken-symmetry solution (see text for further details).

Transitions	Relativistic (eV)	Correlation (eV)
First row 1s	0.2	0.6
Second row 1s	8.0	3.4
Second row 2p	0.5	1.1
All	2.2	1.3

Table 3.6: TiCl_4 pre-edge excitation energies computed with TDDFT and OCDFT using nonrelativistic (NR) and relativistic (DKH2,X2C) Hamiltonians. Results are reported as deviations from the experimental peak maximum (in eV, from ref [121](#)). All calculations use the B3LYP functional and the uncontracted cc-pVDZ basis set.

State	TDDFT		OCDFT		Exp.
	NR	DKH2	NR	X2C	
1	-107.9	-78.6	-35.6	-5.0	} 4969.2
2	-107.9	-78.6	-35.5	-4.5	
3	-107.0	-77.6	-36.4	-5.7	
4	-106.9	-77.6	-35.3	-4.8	
5	-107.0	-77.7	-36.3	-5.7	

minated Ti complexes makes these molecules an ideal test case for investigating the role of relativistic effects on the computation of NEXAS spectra. DeBeer et al. [121](#) obtained the experimental NEXAS spectra of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 , and explained peak positions and intensities using ground state nonrelativistic DFT calculations. Ziegler and co-workers [122](#) studied this class of Ti complexes, employing a relativistic TDDFT approach that made use of the two-component scaled ZORA variant [88](#). While the ZORA-TDDFT spectra displayed excellent qualitative agreement with experimental features [121](#), core excitation energies were underestimated by about 18 eV.

In this work we focus on the analysis of the Ti K-edge spectrum of the three

organotitanium compounds. Table 3.6 reports the pre-edge excitations of TiCl_4 computed with OCDFT and TDDFT using nonrelativistic and relativistic Hamiltonians. These results highlight the significant difference in accuracy between the OCDFT and TDDFT Ti K-edge excitation energies. For example, Ti 1s core excitation energies computed at the nonrelativistic TDDFT level of theory are underestimated, on average, by 107.3 eV compared to the experimental peak position. The addition of scalar relativistic effects via the DKH2 Hamiltonian reduces the TDDFT error to 78.0 eV, which is still a significant deviation from experiment. Nonrelativistic OCDFT core excitation energies underestimate the experimental results by 35.8 eV. Inclusion of scalar relativistic effects via the X2C Hamiltonian reduces this error to 5.1 eV, showing much better agreement with the experimental pre-edge position (4969.2 eV). Note that since we do not have access to a X2C-TDDFT implementation our results are based on the DKH2 Hamiltonian. The comparison between X2C and DKH2 results is justified by the fact that these two methods yield almost identical excitation energies for the test sets. Moreover, both DKH2 and X2C shift the Ti core excitation energy by ca. 30 eV.

The unshifted X2C-OCDFT Ti K-edge peaks of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 are reported in Table 3.7, while Figure 3.4 shows a theoretical spectra simulated using 25 core excited states, together with a comparison of the shifted spectra with the experimental pre-edge features.¹²¹ In accordance with experiment¹²¹, X2C-OCDFT predicts the most intense absorption for TiCl_4 while the least intense for TiCp_2Cl_2 . The relative ordering of the excitation energies of the three complexes predicated by OCDFT is in agreement with experiment as well. The most intense pre-edge feature for TiCl_4 occurs at 4964.4 eV (Expt. 4969.2 eV), while for TiCpCl_3 , and TiCp_2Cl_2 it occurs at 4964.9 (Expt. 4968.1 eV) and 4964.3 (Expt. 4967.3) eV respectively. It is also worth noting that the magnitude of the peak splitting is represented well by OCDFT. Our calculated pre-edge peaks for TiCpCl_3 and TiCp_2Cl_2 are split by 1.4

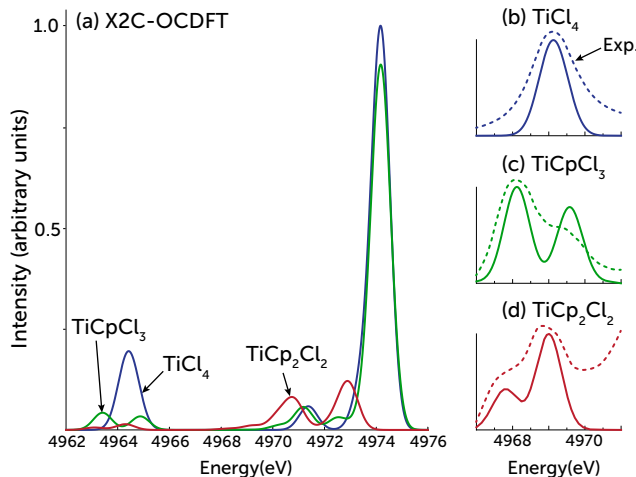


Figure 3.4: (a) Near-edge X-ray absorption spectrum of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 computed with X2C-OCDFT at the B3LYP/un-cc-pVDZ level of theory. (b)–(d) Comparison of the onset of the pre-edge experimental spectra (dotted line) with the X2C-OCDFT spectrum shifted by +4.7 eV to match the experimental pre-edge peak of TiCl_4 (solid line). Experimental curves taken from ref 121. Ti K-edge for each molecule was simulated using 25 unique core-excited state (see supplementary material to see all excited states).

eV and 1.2 eV respectively, which are in excellent agreement with the experimental splitting of 1.4 eV observed for both compounds.

The splitting of the pre-edge peak in the Cp-substituted compounds can be easily understood by examining the hole and particle orbitals that characterize each transition. Figure 3.5 shows the particle orbitals for the five lowest excited states

Table 3.7: Transition energies (in eV) and relative oscillator strengths (f) for pre-edge transitions of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 computed with X2C-OCDFT at the B3LYP/un-cc-pVDZ level of theory. Oscillator strengths shown are values relative to the most intense Ti K-edge transition of these three compounds (see supplementary material for detail).

State	TiCl_4		TiCpCl_3		TiCp_2Cl_2	
	Energy	f	Energy	f	Energy	f
1	4964.2	0.0945	4963.4	0.0202	4963.1	0.0078
2	4964.7	0.0940	4963.4	0.0191	4964.3	0.0121
3	4963.5	0.0002	4963.5	0.0164	4964.4	0.0009
4	4964.4	0.0947	4964.9	0.0220	4964.4	0.0000
5	4963.5	0.0000	4964.9	0.0217	4964.3	0.0055

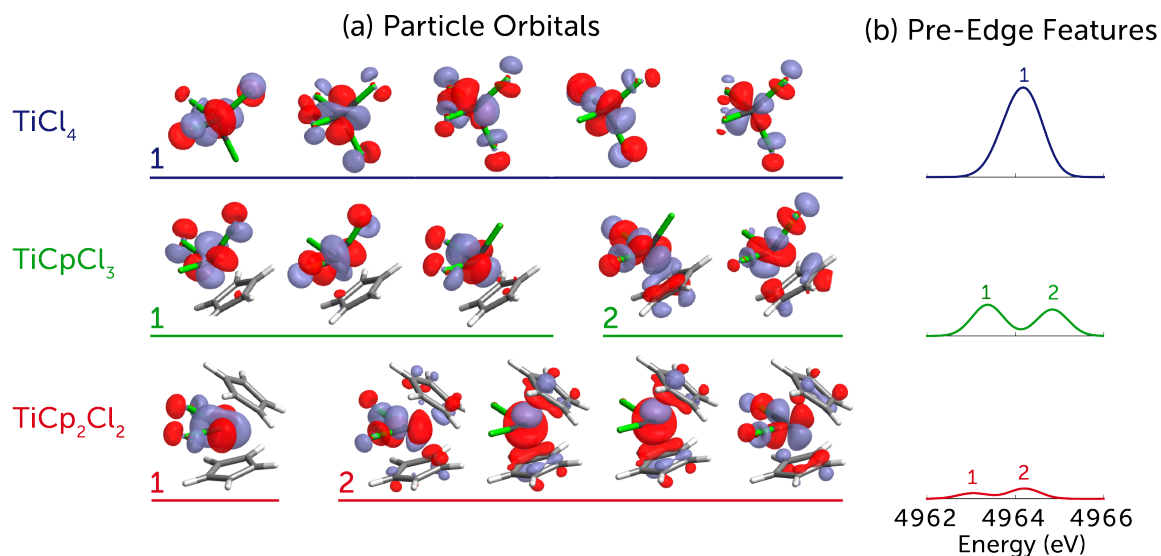


Figure 3.5: (a) Particle orbitals for the core excited states that contribute to the pre-edge features of TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 and (b) corresponding spectral features.

grouped according to which pre-edge feature they contribute. In TiCl_4 all of the particle orbitals are combinations of Ti 3d, Ti 4p, and Cl 3p orbitals. However, upon replacement of Cl with Cp ligands, the particle orbitals split into a group with no Cp ligand contribution (low energy) and a group with significant Cp ligand contribution (high energy). Our assignment of the pre-edge peaks of TiCl_4 and TiCpCl_3 is in agreement with that of Ziegler et al.¹²² However, our assignment of the Ti K-edge of TiCp_2Cl_2 is slightly different from the one reported by Ziegler¹²² and DeBeer.¹²¹ Our OCDFT results suggest that the low energy peak is the result of a single transition while the second peak is a result of two transitions. OCDFT predicts that these three peaks have similar intensities, which results in the 1:2 ratio of the absorption profile. Ziegler et al.¹²² instead find that each peak in the K-edge of TiCp_2Cl_2 is due to a single transition, with the second transition being twice as intense as the first one.

To elucidate the nature of the Ti pre-edge features we generated natural atomic orbitals (NAOs),¹³⁴ for ground and excited states using the JANPA¹³³ software interfaced with PSI4⁹¹. In Table 3.8 we compare the change in orbital shell occupation for each state that contributes to the pre-edge features of the NEXAS spectrum of

Table 3.8: Change in the natural orbital population numbers of the Ti p, Ti d, Cl p shells computed for the Ti pre-edge excitations of TiCl₄, TiCpCl₃, TiCp₂Cl₂.

State	ΔTi_p			ΔCl_p			ΔTi_d		
	TiCl ₄	TiCpCl ₃	TiCp ₂ Cl ₂	TiCl ₄	TiCpCl ₃	TiCp ₂ Cl ₂	TiCl ₄	TiCpCl ₃	TiCp ₂ Cl ₂
1	0.0107	0.0066	0.0097	-0.2184	-0.0597	0.0051	1.1348	1.1614	1.2586
2	0.0126	0.0076	0.0069	-0.3080	-0.0513	-0.0253	1.2111	1.1339	1.1613
3	0.0063	0.0088	0.0048	-0.1790	-0.0437	-0.0920	1.1272	1.1256	1.0685
4	0.0107	0.0055	0.0050	-0.1448	-0.0702	-0.1038	1.0623	1.0368	1.0633
5	0.0064	0.0061	0.0094	-0.0920	-0.0829	-0.0806	1.0403	1.0291	1.0853

TiCl₄, TiCpCl₃, and TiCp₂Cl₂. Our analysis reveals that the pre-edge excitations consist mostly of transitions from Ti s to Ti d orbitals and are accompanied by a small increase in population of the Ti p orbitals, and (on average) a decrease in the population of the Cl p shell.

DeBeer et al.¹²¹ have argued that since the Ti 1s to Ti 3d transition is electric dipole forbidden, the pre-edge states must gain intensity through mixing with metal 4p and ligand 3p orbitals. We have tried to investigate this issue by conducting a Mulliken analysis of the OCDFT transition dipole moment. In OCDFT, the transition dipole moment for the excitation $\Psi^{(0)} \rightarrow \Psi^{(n)}$ is approximated by the matrix element:

$$\boldsymbol{\mu}_{0n} = \langle \Phi^{(n)} | \hat{\mathbf{r}} | \Phi^{(0)} \rangle = \sum_{\mu\nu} D_{\mu\nu}^{0n} \langle \chi_\mu | \hat{\mathbf{r}} | \chi_\nu \rangle \quad (3.11)$$

where $D_{\mu\nu}^{0n}$ is a transition density matrix and $\langle \chi_\mu | \hat{\mathbf{r}} | \chi_\nu \rangle$ is the dipole integral in the atomic orbital basis ($\{\chi_\nu\}$). To analyze each transition we decompose the transition dipole moment into partial atomic contributions. The contribution to the transition dipole moment arising from the angular momentum shell l_A of a donor atom A to the shell l_B on the acceptor atom B is expressed as the restricted sum:

$$\boldsymbol{\mu}_{0n}(A_{l_A} \rightarrow B_{l_B}) = \sum_{\mu \in A_{l_A}} \sum_{\nu \in B_{l_B}} D_{\mu\nu}^{0n} \langle \chi_\mu | \mathbf{r} | \chi_\nu \rangle \quad (3.12)$$

Table 3.9 reports the dominant partial atomic contributions to the transition dipole moment for the pre-edge transitions of TiCpCl₃. For each state we report the three largest contributions to the transition dipole moment, summed over all atoms

of the same type. This analysis reveals that for all pre-edge transitions, the dominant contribution to the transition dipole moment is due to the Ti 4p orbitals. For example, the transition dipole moment of state 1 can be written as a dominant contribution from Ti $s \rightarrow$ Ti p transitions of magnitude $|\boldsymbol{\mu}| = 41.8 \times 10^{-5}$ au, plus two smaller contributions from ligand-ligand and Ti $s \rightarrow$ Cl p transitions with absolute transition dipole moments of 6.6×10^{-5} and 6.4×10^{-5} au, respectively. Other contributions arising from all other combinations of atomic orbitals contribute with a partial transition dipole moment equal to 7.3×10^{-5} au. Note that in many instances the three components of the partial dipole moment have different phases and approximately cancel out. We observe this pattern also for other pre-edge transitions of TiCpCl_3 reported in table 3.9 and for all transitions of TiCl_4 and TiCp_2Cl_2 (see the supplementary material for a full table of all significant atomic contributions to the pre-edge transitions of the three organotitanium complexes). In order to assess the stability of this analysis with respect to the size of the basis set, we computed the $N_{1s} \rightarrow \pi^*$ transition in HCN with six different basis sets: the cc-pVXZ series and the corresponding fully uncontracted basis sets, un-cc-pVXZ (with $X = D, T, Q$). As expected, there is variation in the results, however, in all six cases, the most intense atomic contribution was classified as $N_s \rightarrow N_p$.

Overall, our OCDFt analysis allows us to quantify the contribution of each atom and orbital shell to the NEXAS peak intensity. We found that peak intensity is acquired mainly from Ti $1s \rightarrow$ Ti 4p excitations, a result that is in agreement with that proposed by DeBeer et al.¹²¹ Interestingly, Ti $1s \rightarrow$ Cl 3p excitations produce smaller contributions to the total peak intensity, despite the fact that on average the pre-edge transitions are accompanied by a larger change in the occupation of Cl p orbitals than Ti 4p orbital. This is likely to result from smaller overlap of Ti 1s and Cl p orbitals, which implies a smaller transition dipole moment integral.

Table 3.9: Partial atomic contributions to the transition dipole moment (in atomic units) computed for the pre-edge features of TiCpCl₃.

Contribution	μ_x	μ_y	μ_z	$ \boldsymbol{\mu} $
State 1				
Ti _s → Ti _p	-0.000395	-0.000079	-0.000113	0.000418
Cl _p → Cl _p	+0.000062	+0.000012	+0.000018	0.000066
Ti _s → Cl _p	-0.000006	+0.000003	-0.000064	0.000064
Other	-0.000007	+0.000002	+0.000073	0.000073
State 2				
Ti _s → Ti _p	+0.000119	-0.000321	-0.000228	0.000411
Ti _s → Cl _p	+0.000013	-0.000000	-0.000143	0.000144
Ti _s → C _p	-0.000005	-0.000019	+0.000109	0.000111
Other	-0.000027	+0.000061	+0.000076	0.000101
State 3				
Ti _s → Ti _p	-0.000038	-0.000289	+0.000250	0.000384
Ti _s → Cl _p	-0.000014	-0.000020	+0.000171	0.000173
Ti _s → C _p	+0.000009	-0.000000	-0.000126	0.000126
Other	+0.000016	+0.000057	-0.000098	0.000114
State 4				
Ti _s → Ti _p	-0.000324	-0.000299	+0.000003	0.000441
Cl _p → Cl _p	+0.000135	+0.000128	+0.000005	0.000186
C _p → C _p	-0.000065	-0.000061	-0.000007	0.000089
Other	-0.000021	-0.000015	+0.000001	0.000026
State 5				
Ti _s → Ti _p	+0.000304	-0.000313	+0.000004	0.000436
Cl _p → Cl _p	-0.000128	+0.000134	+0.000003	0.000185
C _p → C _p	+0.000060	-0.000062	+0.000002	0.000086
Other	+0.000021	-0.000024	-0.000007	0.000033

3.5 Discussion and Conclusions

In this study we have combined orthogonality constrained DFT (OCDFT) with the spin-free exact two-component relativistic (X2C) Hamiltonian, and used this as a tool to analyze scalar relativistic effects in the context of core electron excitations. When relativistic effects are not significant, OCDFT consistently reproduces experimental core-valence excitation energies within 1 eV. The influence of scalar relativistic effects can already be seen in the failure of nonrelativistic OCDFT at computing core excitation energies of transitions involving the 1s orbital of second-row elements. For

example, the magnitude of relativistic effects for the Si 1s $\rightarrow \sigma^*$ excitation in SiH₄ is estimated to be 4.3 eV, but for the Cl 1s $\rightarrow \sigma^*$ excitation in HCl this figure increases to 10.1 eV. We demonstrated that by combining OCDFT with a scalar relativistic Hamiltonian and the B3LYP functional, core excitation energies for second row elements can be computed with an average error of 2.3 eV (from a test set of 10 transitions), a substantial improvement over the 10.3 eV average error observed with the nonrelativistic OCDFT.

Using a class of well studied organotitanium compounds (TiCp_nCl_{4-n}, n = 2-4), we have demonstrated the utility of the X2C-OCDFT method as a tool to compute accurate transition metal K-edge spectra. The inclusion of scalar relativistic effects is mandatory to obtain near-quantitative agreement with experimental excitation energies. For example, for TiCpCl₃ we find that a non-relativistic OCDFT treatment predicts excitation energies with an error of about 35 eV. Once augmented with X2C, the OCDFT error is reduced to 4.8 eV, which is a marked improvement results from relativistic TDDFT (18 eV)¹²². We find that X2C-OCDFT consistently underestimates the pre-edge features of Ti 1s transitions by ca. 5 eV, which is commensurate with our analysis of second row 1s excitations (MAE = 2.3 eV, max error = 3.6 eV). In other words, the relative error of X2C-OCDFT Ti K-edge excitation energies is of the order of 0.1%. A large portion of the remaining error is most likely due to the use of an approximate functional for the excited state employed in the theory (adiabatic approximation), but also the increased importance of non-scalar relativistic effects (spin-orbit coupling, finite nuclei effects, etc.) that are not captured by spin-free scalar relativistic Hamiltonians.

To gain insight into the nature of the Ti K-edge excitations, we characterize the change in natural atomic orbital population and examine partial atomic contributions to the transition dipole moment of all states that contribute to the pre-edge features. The character of the pre-edge transitions of the three organotitanium com-

pounds as predicted by OCDFT is consistent with previous analyses,^{121,122} however, in this study we developed an approach to quantify partial atomic contributions to the transition dipole moments. This tool enables us to attribute most of the pre-edge peak intensity to excitations involving Ti 4p orbitals. Ligand orbitals are found to give small contributions to the peak intensity.

What contributes to the improved performance of OCDFT with respect to TDDFT for core excitation energies? Our evidence suggests that like other Δ SCF based approaches,^{16,34,44,47,48,135} OCDFT derives an advantage from the fact that it can properly describe charge-transfer excitations.^{97,98} To illustrate this point, we computed the C and O K-edge spectrum of CO using various functionals and interpreted our data using a toy model of that captures the charge-transfer character of core excitations. Our analysis suggests that a large source of the error in TDDFT arises from double counting the change in Coulomb energy that results from unpairing a core electron. In the case of OCDFT, the effects of double counting the Coulomb repulsion is smaller, and the energy expression involves only local self-interaction terms. We also note a stronger dependence of the TDDFT spectrum on the amount of Hartree–Fock exchange included in the functional. Although shifted TDDFT will be comparable to OCDFT results, they are not identical, and there exists significant merit in performing OCDFT calculations since it yields more consistent predictions.

One future application of this work is to extend OCDFT to the computation of L-edge spectra. The L-edge is composed of excitations from core 2p orbitals and simulation of the full spectrum will require proper treatment of spin-orbit coupling effects due to mixing of excitations from degenerate 2p orbitals. We would also like to explore the idea of using this orthogonality constrained excited state framework to build wave function based methods with the same advantages seen in OCDFT. This would allow us to create systematically improvable schemes that do not rely on approximate exchange-correlation functionals.

Acknowledgments

This work was supported by start-up funds provided by Emory University. We would like to thank Dr. Lan Cheng for providing data to benchmark our implementation of X2C. We would also like to acknowledge Dr. Tymofii Y. Nikolaienko for providing help with the JANPA software package. W.D.D. is supported by the National Science Foundation Graduate Research Fellowship under Grant No. 0000048655. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Bibliography

- [1] Stöhr, J. *NEXAFS Spectroscopy*; Springer Series in Surface Sciences; Springer, 1992.
- [2] Westre, T. E.; Kennepohl, P.; DeWitt, J. G.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1997**, *119*, 6297–6314.
- [3] Löble, M. W. et al. *J. Am. Chem. Soc.* **2015**, *137*, 2506–2523.
- [4] Idrees, M.; Nadeem, M.; Sung, N.-E.; Asanova, T.; Shin, T. J. *Chem. Phys. Lett.* **2014**, *612*, 262–265.
- [5] Nelson, R. C.; Miller, J. T. *Catal. Sci. Technol.* **2012**, *2*, 461–470.
- [6] Contini, G.; Carravetta, V.; Di Castro, V.; Stranges, S.; Richter, R.; Alagia, M. *J. Phys. Chem. A* **2001**, *105*, 7308–7314.
- [7] Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Näslund, L.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995–999.
- [8] Freiwald, M.; Cramm, S.; Eberhardt, W.; Eisebitt, S. *J. Electron Spectrosc. Relat. Phenom.* **2004**, *137–140*, 413–416.
- [9] Hua, W.; Gao, B.; Li, S.; Ågren, H.; Luo, Y. *J. Phys. Chem. B* **2010**, *114*, 13214–13222.
- [10] Hahner, G. *Chem. Soc. Rev.* **2006**, *35*, 1244–1255.
- [11] Guo, Q. X.; Senda, H.; Saito, K.; Tanaka, T.; Nishio, M.; Ding, J.; Fan, T. X.; Zhang, D.; Wang, X. Q.; Liu, S. T.; Shen, B.; Ohtani, R. *Appl. Phys. Lett.* **2011**, *98*, 181901.
- [12] Milne, C.; Penfold, T.; Chergui, M. *Coord. Chem. Rev.* **2014**, *277–278*, 44–68.
- [13] Cabaret, D.; Bordage, A.; Juhin, A.; Arfaoui, M.; Gaudry, E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5619–5633.

- [14] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A* **2012**, *85*, 022507.
- [15] Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 6735–6756.
- [16] Besley, N. A. *Chem. Phys. Lett.* **2012**, *542*, 42–46.
- [17] Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.
- [18] Dutta, A. K.; Gupta, J.; Vaval, N.; Pal, S. *J. Chem. Theory Comput.* **2014**, *10*, 3656–3668.
- [19] Brabec, J.; Bhaskaran-Nair, K.; Govind, N.; Pittner, J.; Kowalski, K. *J. Chem. Phys.* **2012**, *137*, 171101.
- [20] Sen, S.; Shee, A.; Mukherjee, D. *Mol. Phys.* **2013**, *111*, 2625–2639.
- [21] Barth, A.; Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. *Chem. Phys.* **1980**, *46*, 149–164.
- [22] Butscher, W.; Buenker, R. J.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1977**, *52*, 449–456.
- [23] Roemelt, M.; Neese, F. *J. Phys. Chem. A* **2013**, *117*, 3069–3083.
- [24] Asmuruf, F.; Besley, N. *Chem. Phys. Lett.* **2008**, *463*, 267–271.
- [25] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.
- [26] Schirmer, J. *Phys. Rev. A* **1982**, *26*, 2395–2416.
- [27] Salpeter, E. E.; Bethe, H. A. *Phys. Rev.* **1951**, *84*, 1232–1242.
- [28] Rehr, J. J.; Albers, R. C. *Rev. Mod. Phys.* **2000**, *72*, 621–654.
- [29] Natoli, C. R.; Krüger, P.; Hatada, K.; Hayakawa, K.; Sébilleau, D.; Šipr, O. *J. Phys-Condens. Mat.* **2012**, *24*, 365501.
- [30] Joly, Y. *Phys. Rev. B* **2001**, *63*, 125120.
- [31] Slater, J. C. In *Statistical Exchange-Correlation in the Self-Consistent Field*; Löwdin, P.-O., Ed.; Adv. Quantum Chem.; Academic Press, 1972; Vol. 6; pp 1–92.
- [32] Triguero, L.; Pettersson, L. G. M.; Ågren, H. *Phys. Rev. B* **1998**, *58*, 8097–8110.
- [33] Prendergast, D.; Galli, G. *Phys. Rev. Lett.* **2006**, *96*, 215502.
- [34] Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. *Chem. Phys. Lett.* **1994**, *222*, 75–81.
- [35] Kuramoto, K.; Ehara, M.; Nakatsuji, H. *J. Chem. Phys.* **2005**, *122*, 014304.
- [36] Ohtsuka, Y.; Nakatsuji, H. *J. Chem. Phys.* **2006**, *124*, 054110.
- [37] Kolczewski, C.; Püttner, R.; Plashkevych, O.; Ågren, H.; Staemmler, V.; Martins, M.; Snell, G.; Schlachter, A. S.; Sant’Anna, M.; Kaindl, G.; Pettersson, L. G. M. *J. Chem. Phys.* **2001**, *115*, 6426–6437.
- [38] Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- [39] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039.
- [40] Lestrange, P. J.; Nguyen, P. D.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.

- [41] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.
- [42] Fernando, R. G.; Balhoff, M. C.; Lopata, K. *J. Chem. Theory Comput.* **2015**, *11*, 646–654.
- [43] Casida, M.; Huix-Rotllant, M. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.
- [44] Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. *J. Chem. Phys.* **2009**, *130*, 124308.
- [45] Verma, P.; Bartlett, R. J. *J. Chem. Phys.* **2012**, *137*, 134102.
- [46] Verma, P.; Bartlett, R. J. *J. Chem. Phys.* **2014**, *140*, 18A534.
- [47] Kowalczyk, T.; Yost, S. R.; Van Voorhis, T. *J. Chem. Phys.* **2011**, *134*, 054128.
- [48] Krykunov, M.; Ziegler, T. *J. Chem. Theory Comput.* **2013**, *9*, 2761–2773.
- [49] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- [50] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, 14360–14374.
- [51] Dyall, K. G.; Faegri, K. *Introduction to Relativistic Quantum Chemistry*; Oxford University Press, New York, 2007.
- [52] Reiher, M.; Wolf, A. *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science*; Wiley-VCH: Weinheim, Germany, 2009.
- [53] Moss, R. *Advanced Molecular Quantum Mechanics*; Studies in chemical physics; Chapman and Hall, London, 1973.
- [54] Grant, I. *Relativistic Quantum Theory of Atoms and Molecules: Theory and Computation*; Springer Series on Atomic, Optical, and Plasma Physics; Springer-Verlag, New York, 2007.
- [55] Kutzelnigg, W. *Chem. Phys.* **2012**, *395*, 16–34.
- [56] Cremer, D.; Zou, W.; Filatov, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 436–467.
- [57] Peng, D.; Reiher, M. *Theor. Chem. Acc.* **2012**, *131*, 1081.
- [58] Autschbach, J. *J. Chem. Phys.* **2012**, *136*, 150902.
- [59] Saue, T. *ChemPhysChem* **2011**, *12*, 3077–3094.
- [60] Liu, W. *Mol. Phys.* **2010**, *108*, 1679–1706.
- [61] Schwerdtfeger, P. *Relativistic Electronic Structure Theory: Part 2. Applications*; Theoretical and Computational Chemistry; Elsevier Science, 2004.
- [62] Pyykkö, P. *Chem. Rev.* **1988**, *88*, 563–594.
- [63] Foldy, L. L.; Wouthuysen, S. A. *Phys. Rev.* **1950**, *78*, 29–36.
- [64] Heully, J. L.; Lindgren, I.; Lindroth, E.; Lundqvist, S.; Martensson-Pendrill, A. M. *J. Phys. B* **1986**, *19*, 2799.
- [65] Chang, C.; Pelissier, M.; Durand, P. *Phys. Scr.* **1986**, *34*, 394.
- [66] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1993**, *99*, 4597–4610.

- [67] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1994**, *101*, 9783–9792.
- [68] Reiher, M.; Wolf, A. *J. Chem. Phys.* **2004**, *121*, 10945–10956.
- [69] Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89 – 155.
- [70] van Wüllen, C. *J. Chem. Phys.* **2004**, *120*, 7307–7313.
- [71] Wolf, A.; Reiher, M.; Hess, B. A. *J. Chem. Phys.* **2002**, *117*, 9215–9226.
- [72] Nakajima, T.; Hirao, K. *J. Chem. Phys.* **2000**, *113*, 7786–7789.
- [73] Dyal, K. G. *J. Chem. Phys.* **1997**, *106*, 9618–9626.
- [74] Dyal, K. *J. Chem. Phys.* **2001**, *115*, 9136–9143.
- [75] Filatov, M.; Dyal, K. *Theor. Chem. Acc.* **2007**, *117*, 333–338.
- [76] Zou, W.; Filatov, M.; Cremer, D. *Theor. Chem. Acc.* **2011**, *130*, 633–644.
- [77] Barysz, M.; Sadlej, A. J.; Snijders, J. G. *Int. J. Quantum Chem.* **1997**, *65*, 225–239.
- [78] Barysz, M.; Sadlej, A. J. *J. Chem. Phys.* **2002**, *116*, 2696–2704.
- [79] Barysz, M.; Mentel, L.; Leszczyński, J. *J. Chem. Phys.* **2009**, *130*, 164114.
- [80] Kutzelnigg, W.; Liu, W. *J. Chem. Phys.* **2005**, *123*, 241102.
- [81] Iliáš, M.; Jensen, H. J. A.; Kelló, V.; Roos, B. O.; Urban, M. *Chem. Phys. Lett.* **2005**, *408*, 210–215.
- [82] Liu, W.; Peng, D. *J. Chem. Phys.* **2006**, *125*, 044102.
- [83] Liu, W.; Kutzelnigg, W. *J. Chem. Phys.* **2007**, *126*, 114107.
- [84] Iliáš, M.; Saue, T. *J. Chem. Phys.* **2007**, *126*, 064102.
- [85] Peng, D.; Liu, W.; Xiao, Y.; Cheng, L. *J. Chem. Phys.* **2007**, *127*, 104106.
- [86] Liu, W.; Peng, D. *J. Chem. Phys.* **2009**, *131*, 031104.
- [87] van Wüllen, C. *J. Chem. Phys.* **1998**, *109*, 392–399.
- [88] Philipsen, P. H. T.; van Lenthe, E.; Snijders, J. G.; Baerends, E. J. *Phys. Rev. B* **1997**, *56*, 13556–13562.
- [89] Nichols, P.; Govind, N.; Bylaska, E. J.; de Jong, W. A. *J. Chem. Theory Comput.* **2009**, *5*, 491–499.
- [90] van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1994**, *101*, 9783–9792.
- [91] Turney, J. M. et al. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 556–565.
- [92] Kutzelnigg, W. *Int. J. Quantum Chem.* **1984**, *25*, 107–129.
- [93] Dyal, K. G.; Grant, I. P.; Wilson, S. *J. Phys. B* **1984**, *17*, 493.
- [94] Cheng, L.; Gauss, J. *J. Chem. Phys.* **2011**, *135*, 084114.

- [95] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A* **2012**, *85*, 042518.
- [96] Besley, N. A.; Noble, A. *J. Phys. Chem. C* **2007**, *111*, 3333–3340.
- [97] Casida, M. E.; Gutierrez, F.; Guan, J.; Gadea, F.-X.; Salahub, D.; Daudey, J.-P. *J. Chem. Phys.* **2000**, *113*, 7062–7071.
- [98] Ziegler, T.; Krykunov, M.; Cullen, J. *J. Chem. Phys.* **2012**, *136*, 124107.
- [99] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- [100] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [101] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [102] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [103] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [104] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- [105] Püttner, R.; Dominguez, I.; Morgan, T. J.; Cisneros, C.; Fink, R. F.; Rotenberg, E.; Warwick, T.; Domke, M.; Kaindl, G.; Schlachter, A. S. *Phys. Rev. A* **1999**, *59*, 3415.
- [106] Remmers, G.; Domke, M.; Puschmann, A.; Mandel, T.; Xue, C.; Kaindl, G.; Hudson, E.; Shirley, D. A. *Phys. Rev. A* **1992**, *46*, 3935.
- [107] Domke, M.; Xue, C.; Puschmann, A.; Mandel, T.; Hudson, E.; Shirley, D.; Kaindl, G. *Chem. Phys. Lett.* **1990**, *173*, 122–128.
- [108] Chen, C. T.; Ma, Y.; Sette, F. *Phys. Rev. A* **1989**, *40*, 6737.
- [109] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1980**, *13*, 999.
- [110] Tronc, M.; King, G. C.; Read, F. H. *J. Phys. B: At., Mol. Phys.* **1979**, *12*, 137.
- [111] Francis, J. T.; Enkvist, C.; Lunell, S.; Hitchcock, A. P. *Can. J. Phys.* **1994**, *72*, 879–884.
- [112] Adachi, J.-i.; Kosugi, N.; Shigemasa, E.; Yagishita, A. *Chem. Phys. Lett.* **1999**, *309*, 427–433.
- [113] Hitchcock, A. P.; Brion, C. E. *J. Electron Spectrosc. Relat. Phenom.* **1979**, *15*, 201–206.
- [114] Domke, M.; Xue, C.; Puschmann, A.; Mandel, T.; Hudson, E.; Shirley, D. A.; Kaindl, G. *Chem. Phys. Lett.* **1990**, *173*, 122–128.
- [115] Nayandin, O.; Kukk, E.; Wills, A.; Langer, B.; Bozek, J.; Canton-Rogan, S.; Wiedenhoeft, M.; Cubaynes, D.; Berrah, N. *Phys. Rev. A* **2001**, *63*, 062719.
- [116] Bodeur, S.; Millié, P.; Nenner, I. *Phys. Rev. A* **1990**, *41*, 252.
- [117] Gedat, E.; Püttner, R.; Domke, M.; Kaindl, G. *J. Chem. Phys.* **1998**, *109*, 4471.
- [118] Hudson, E.; Shirley, D. A.; Domke, M.; Remmers, G.; Kaindl, G. *Phys. Rev. A* **1994**, *49*, 161–175.
- [119] Cavell, R. G.; Jürgensen, A. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101–103*, 125–129.
- [120] Bodeur, S.; Esteva, J. M. *Chem. Phys.* **1985**, *100*, 415–427.

- [121] DeBeer George, S.; Brant, P.; Solomon, E. I. *J. Am. Chem. Soc.* **2005**, *127*, 667–674.
- [122] Casarin, M.; Finetti, P.; Vittadini, A.; Wang, F.; Ziegler, T. *J. Phys. Chem. A* **2007**, *111*, 5270–5279.
- [123] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- [124] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [125] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993–4999.
- [126] Schrader, D. M.; Prager, S. *J. Chem. Phys.* **1962**, *37*, 1456–1460.
- [127] Baerends, E.; Ellis, D.; Ros, P. *Chem. Phys.* **1973**, *2*, 41 – 51.
- [128] Dunlap, B. I. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2113–2116.
- [129] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *Int. J. Quantum Chem., Symp.* **1977**, *12*, 81–87.
- [130] Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.
- [131] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- [132] Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- [133] Nikolaienko, T. Y.; Bulavin, L. A.; Hovorun, D. M. *Comp. Theor. Chem.* **2014**, *1050*, 15–22.
- [134] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
- [135] Fransson, T.; Coriani, S.; Christiansen, O.; Norman, P. *J. Chem. Phys.* **2013**, *138*, 124311.

Chapter 4

Localized Intrinsic Valence Virtual Orbitals as a Tool for the Automated Classification of Core Excited States

“We must be clear that when it comes to atoms, language can be used as in poetry. The poet, too, is not nearly so concerned with describing facts as with creating images and establishing mental connections”

Niels H. D. Bohr

Chapter Abstract

Accurate assignments of the unoccupied molecular orbitals involved in electronic excited states are crucial to the interpretation of experimental spectra. Here we present an automated approach to the assignment of excited states by introducing a unique orbital basis known as localized intrinsic valence virtual orbitals (LIVVOs), which are a special case of the previously reported valence virtual orbitals. The LIVVOs are used to quantify the local contributions to particle orbitals from orthogonality-constrained density functional theory, providing an assignment with atomic-level/angular momentum shell specificity. This localized set also allows us to define the total valence character of an excited state. We highlight the utility of our approach by studying the local orbital changes in core-excited states at the sulfur K-edge of ethanethiol and benzenethiol as well as hydrogen bonding in the water dimer.

4.1 Introduction

Any interpretation of electronic spectroscopy depends upon assigning features of the experimental spectrum to discrete electronic transitions calculated using excited state quantum chemistry methods. This requires characterization of the molecular orbitals involved in the discrete transition. Transitions are usually assigned by visual inspection of the contributing virtual molecular orbitals (MOs), which requires that the orbitals have a sensible shape such that the orbital character is comprehensible to the user. Unfortunately, the canonical virtual MOs produced in quantum chemistry calculations produce a large and diverse set where the virtual orbitals that are useful for describing excitation processes exist within a continuum of states that describe detached electrons. Often, the MOs within this large set do not have well defined orbital character, orbital labels are used instead (e.g. LUMO, LUMO + 1, etc.). These approaches are somewhat unsatisfactory. The assignment of orbital character based on MO plots is potentially ambiguous. Furthermore, characterizing a transition using orbital labels is problematic because virtual orbitals are not uniquely defined and they strongly depend on the basis set and level of theory employed.

A unique branch of electronic spectroscopy involving the excitation of core electrons using X-ray photons called near-edge X-ray absorption fine structure (NEXAFS) spectroscopy has garnered particular interest due to advancements in high-resolution tunable synchrotron light sources. NEXAFS spectroscopy has become an indispensable experimental tool for probing local electronic and geometrical information in a wide variety of molecular systems.¹⁻⁴ The significant features of the NEXAFS spectrum arise from core orbital to unoccupied valence orbital transitions.⁵ The atomic nature of the core orbitals involved in these transitions (1s, 2s, 2p, etc.) makes NEXAFS a spectroscopic technique for elucidating local electronic structure information. Quantum chemistry calculations serve as a vital tool for gleaning meaningful

electronic structure information from NEXAFS spectra including the nature of unoccupied orbitals, orbital mixing effects, and covalent character.^{2,5-8} A myriad of theoretical methods have been developed for the computation of core-excited states, ranging from rigorous methods focused on high accuracy⁹⁻²⁴ to computationally efficient methods focused on applications to larger systems.²⁵⁻³⁴ The primary motivation underpinning these theoretical developments has been to obtain more accurate excitation energies and/or oscillator strengths. In contrast, the character of the wave function is analyzed in a more cursory fashion.

A recent work by the group of Dreuw³⁵⁻³⁷ has taken a significant step toward a more rigorous description of core-excited states. Using the one-particle transition density matrix and the one-particle difference density matrix, these authors were able to quantify the effects of relaxation, estimate multireference character, and provide information on the exciton size. These methods are a significant step toward a robust characterization of excited-states. Myhre, Coriani, and Koch³⁸ also recently used visualization of the one-particle difference density matrix in order to track the localization of core-excited states, relying on analysis of the excitation amplitudes of coupled cluster response calculations to assign the character of core-valence transitions. Although a fair amount of attention has been put forth to add a breadth of different descriptors for excited states, little attention has been put toward trying to improve upon the method of assigning orbital character.

Canonical molecular orbitals often complicate the assignment process because valence virtual MOs are mixed with polarization and diffuse functions included in the basis set. A common approach to obtain more chemically meaningful valence orbitals is to project the virtual space onto a set of atomic valence orbitals. This approach hinges on the assumption that valence atomic orbitals are the dominant contributors to the valence molecular orbitals. This assumption was justified in the 1980s by Rudenberg and coworkers by obtaining overlaps close to unity between

free-atom AOs and their projection onto the molecular orbitals of MCSCF wave functions.³⁹ Perhaps, the most popular method to analyze wave functions in terms of atomic valence contributions is Weinhold’s natural atomic/bond orbital (NAO/NBO) analysis.^{40,41} This method partitions the full density matrix into atomic sub-blocks that are diagonalized in order to produce atom-localized hybrid valence orbitals. AO projection based methods have been presented throughout the literature including polarized atomic orbitals,^{42,43} enveloping localized orbitals,⁴⁴ intrinsic minimal atomic basis orbitals,⁴⁵ molecule-adapted atomic orbitals,⁴⁶ and intrinsic atomic orbitals.⁴⁷ The approach presented in this work is based on the valence virtual orbitals (VVOs) scheme.⁴⁸ VVOs are built by bringing virtual valence MOs into maximum coincidence with a minimal atomic basis set via a unitary transformation. The VVOs have been used with success to identify the valence LUMO in a variety of molecular systems,⁴⁹ to characterize bond breaking/formation along dissociation pathways,⁵⁰ and as a set of starting orbitals for multi-configurational SCF calculations.⁵¹ VVOs can be easily localized and provide an accurate representation of the localized valence antibonding orbitals in a molecule.

In this work we report a new method for automated classification of virtual orbitals in electronic excited states by quantifying the local contributions using a localized set of VVOs. To generate the VVOs an accurate atomic minimal basis set (AAMBS) that spans the atomic valence space is required. We propose to represent the AAMBS using a minimal basis set of quasiautomatic orbitals.⁴⁸ These are generated using projection operators according to Knizia’s intrinsic atomic orbitals (IAOs) procedure⁴⁷ and localized using the Pipek–Mezey localization scheme.⁵² The combination of VVOs projected onto a localized intrinsic orbital basis is abbreviated as LIVVOs. To compute excitation energies we employ our recently proposed orthogonality-constrained DFT (OCDFT) approach⁵³ to compute a manifold of states.⁵⁴ OCDFT is a variational time-independent DFT method built upon a generalized Kohn-Sham scheme.

In OCDFT excited state energies are approximated by a Kohn–Sham procedure augmented with orthogonality constraints between ground and excited state wave functions.

This new classification method is applied to the core-excited states involved in NEXAFS spectroscopy. We utilize the quantification of local contributions to analyze substituent effects in thiols (ethanethiol, benzenethiol) and intermolecular interactions in the water dimer. The thiols were chosen as an example of molecules where the relevant excited states have significant contributions from more than one valence antibonding MO. Previous studies have not fully characterized the antibonding character of orbitals involved in core-valence excitations.^{55–57} We show that we are able to provide insight into the nature of these excited states using LIVVOs. Hydrogen bonding interactions in water have received plenty of attention within the X-ray absorption community.^{58–64} Studies have shown that NEXAFS spectral changes upon the formation of a hydrogen bond is a result of localization of excited states along either the free OH bond or the hydrogen bond.^{60,61,64–70} However, the degree of localization has not been quantified. In this work, we show that LIVVOs help quantify the localization of core excited states of the water dimer and that this metric is correlated to some spectral features.

4.2 Theory

In this section we discuss three aspects of our approach to analyze core excitations. First, we consider the construction of the localized intrinsic valence virtual orbitals (LIVVOs) from canonical molecular orbitals (CMOs) and intrinsic atomic orbitals (IAOs). Next, we discuss how to automate the determination of the atomic character of the IAOs and the LIVVOs. Finally, we discuss the procedure used to analyze the particle orbital in OCDFT computations of core-excited states.

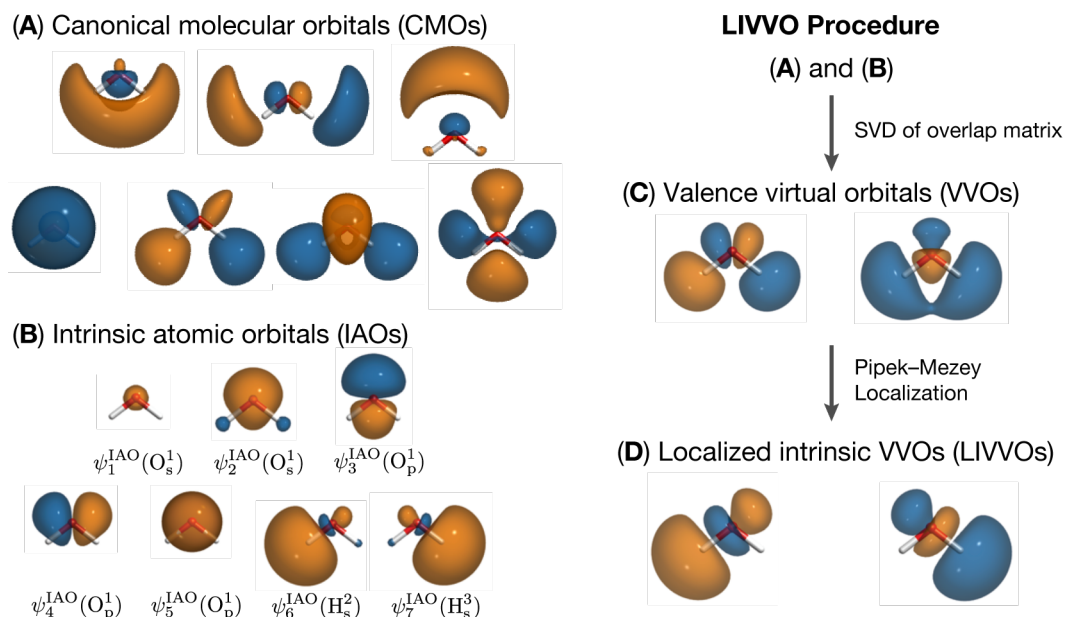


Figure 4.1: Illustration of the procedure used to construct localized intrinsic valence virtual orbitals (LIVVOs). (A) Canonical molecular orbitals (CMOs) of water computed at the B3LYP/aug-cc-pVTZ level of theory. (B) Intrinsic atomic orbitals obtained from the CMOs of water and corresponding assignment to a given atom and angular momentum shell. (C) Delocalized valence virtual orbitals (VVOs) obtained by singular value decomposition of the overlap matrix [eq (4.3)]. (D) Localized intrinsic valence virtual orbitals (LIVVOs) for water after Pipek-Mezey localization of the VVOs.

4.2.1 Construction of Localized Intrinsic Valence Virtual Orbitals

Here we summarize the construction of VVOs and discuss how to obtain the LIVVOs. The procedure outlined here is illustrated in Figure 4.1 for the case of water. The construction of VVOs begins by first identifying the virtual subset of the canonical molecular orbitals ($\{\phi_a\}$) from a ground state SCF calculation. The virtual CMOs are expanded in an atom-centered basis $\{\chi_\mu\}$ and are expressed in terms of MO coefficients $C_{\mu a}$

$$|\phi_a\rangle = \sum_{\mu=1}^{N_{\text{AO}}} |\chi_\mu\rangle C_{\mu a}, \quad a = 1, \dots, N_{\text{vir}} \quad (4.1)$$

where N_{AO} and N_{vir} are the number of atomic and virtual orbitals, respectively. As shown in Figure 4.1A, the virtual CMOs may have contributions from diffuse functions

and their valence character is difficult to interpret.

To identify the valence character of the CMOs, we perform a VVO analysis^{49,71,72} using the intrinsic atomic orbitals (IAOs) of Knizia⁴⁷ as the underlying accurate atomic minimal basis set. IAOs are a set of orthonormal polarized atomic orbitals that can exactly express the occupied molecular orbitals of the ground state Kohn-Sham (KS) determinant $|\Phi\rangle$. IAOs are expressed in terms of the atomic orbital basis as

$$|\psi_\rho^{\text{IAO}}\rangle = \sum_{\mu=1}^{N_{\text{AO}}} |\chi_\mu\rangle \tilde{C}_{\mu\rho}, \quad \rho = 1, \dots, N_{\text{IAO}} \quad (4.2)$$

where $\tilde{C}_{\mu\rho}$ is the IAO coefficient matrix and N_{IAO} is the number of IAO orbitals. As shown in Figure 4.1B, in the case of water there are 7 IAOs that correspond to the oxygen atom 1s, 2s, and 2p shells and a single 1s orbital for each hydrogen atom.

Next, the overlap matrix (\mathbf{S}) of the canonical virtual orbitals with the IAO functions ψ_ρ^{IAO} is evaluated:

$$(\mathbf{S})_{a\rho} = \langle \phi_a | \psi_\rho^{\text{IAO}} \rangle \quad (4.3)$$

and as suggested in Ref. 49 a singular value decomposition (SVD) is performed on \mathbf{S} :

$$\mathbf{S} = \mathbf{U} \boldsymbol{\sigma} \mathbf{V}^\dagger \quad (4.4)$$

to yield orthogonal transformation matrices \mathbf{U} and \mathbf{V} . These matrices are rotations of the virtual space and the IAO space, respectively, that bring the two sets of orbitals into maximum coincidence. The number of VVOs (N_{VVO}) is given by the number of IAOs minus the number of occupied molecular (N_{occ}) orbitals: $N_{\text{VVO}} = N_{\text{IAO}} - N_{\text{occ}}$. Assuming that the singular value decomposition orders the singular values (σ_v) in descending order, then the VVOs are obtained by transforming the canonical virtual orbitals with the first N_{VVO} columns of the matrix \mathbf{U} :

$$|\psi_v^{\text{VVO}}\rangle = \sum_{a=1}^{N_{\text{vir}}} |\phi_a\rangle U_{av}, \quad v = 1, \dots, N_{\text{VVO}}. \quad (4.5)$$

The VVOs for water shown in Figure 4.1C, have valence character and correspond to the bonding and antibonding combination of localized $\sigma_{\text{O-H}}^*$ orbitals. Following Pipek–Mezey localization,⁵² we arrive at a set of localized VVOs that span the antibonding interactions of the molecular environment. We refer to this localized set as localized intrinsic valence virtual orbitals (LIVVOs). In the case of water, following localization of the VVOs we obtain the $\sigma_{\text{O-H}}^*$ orbitals drawn in Figure 4.1D.

4.2.2 Determination of the character of the IAOs and LIVVOs

The goal of our analysis is to express excited states using the LIVVO basis. In addition, we are also interested in determining the dominant atomic character of each excited state. To this end, we have automated the analysis of the atomic character of the LIVVOs. Our approach starts with the assignment of the atomic character of each IAO via a Mulliken population analysis.^{73,74} For each IAO, ψ_{ρ}^{IAO} , we compute the population matrix:

$$M_{\mu\nu}(\rho) = \tilde{C}_{\mu\rho} S_{\mu\nu} \tilde{C}_{\nu\rho}, \quad \rho = 1, \dots, N_{\text{IAO}} \quad (4.6)$$

Gross populations on atom A and angular momentum shell l [GP_{A_l}] are obtained as partial sums of the population matrix:

$$\text{GP}_{A_l} = \sum_{\mu \in A_l} \sum_{\nu=1}^{N_{\text{IAO}}} M_{\mu\nu}(\rho) \quad (4.7)$$

where the sum over μ is restricted to all functions centered on atom A with angular momentum quantum number l .

Values of GP_{A_l} for the water example are reported in Table 4.1. The assignment of IAOs using the gross population is straightforward since there is always a dominant contribution from a single atom/shell. For example, in the case of IAO ψ_6^{IAO} , the largest contribution to GP_{A_l} is from the H_s^2 shell (1.18), followed by smaller contributions from the O_{1s}^1 , O_{2s}^1 , O_{2p}^1 , and H_s^3 shells (with absolute value less or equal

to 0.08). Therefore, we define the character of ψ_ρ^{IAO} as the pair atom/shell (A,l) for which GP_{A_l} is maximum:

$$\text{character}(\psi_\rho^{\text{IAO}}) = \arg \max_{(A,l)} \text{GP}_{A_l} \quad (4.8)$$

In the case of water, when eq (4.8) is applied to the IAOs we obtain the assignment reported in Figure 4.1B.

Once the atomic character of each IAO is determined, we characterize the atomic contributions to each LIVVO. For each LIVVO, ψ_l^{LIVVO} , we evaluate the overlap with all the IAOs (ψ_ρ^{IAO}):

$$S'_{l\rho} = |\langle \psi_l^{\text{LIVVO}} | \psi_\rho^{\text{IAO}} \rangle|^2 \quad (4.9)$$

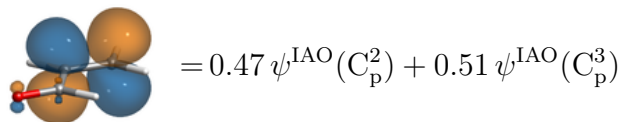
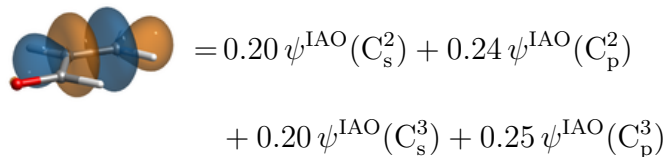
The initial step in classifying the LIVVOs is to determine their overall orbital character (σ, π, \dots). A LIVVO is classified as σ^* if the largest elements of $S'_{l\rho}$ include contributions from IAOs with s and p character. Similarly, a LIVVO is assigned π^* character if the largest elements of $S'_{l\rho}$ arise exclusively from p -type IAOs. The atoms involved in the LIVVO are identified by the character of the largest elements of $S'_{l\rho}$. For this purpose, we only include those IAOs whose overlap with a given LIVVO is greater than a threshold, which in the following examples is set to 0.1.

To illustrate this analysis we consider the case of acrolein ($\text{C}_3\text{H}_4\text{O}$). The following

Table 4.1: Atom-centered gross orbital populations $[\text{GP}_{A_l}]$ for the seven IAOs of water. Calculations were performed at the B3LYP/jun-cc-pVTZ level of theory.

A_l	ψ_1^{IAO}	ψ_2^{IAO}	ψ_3^{IAO}	ψ_4^{IAO}	ψ_5^{IAO}	ψ_6^{IAO}	ψ_7^{IAO}
O_s^1	1.00	1.16	0.00	0.00	0.00	-0.08	-0.08
O_p^1	0.00	0.00	1.04	1.11	1.00	-0.07	-0.07
H_s^2	0.00	-0.08	-0.02	-0.05	0.00	1.18	-0.03
H_s^3	0.00	-0.08	-0.02	-0.05	0.00	-0.03	1.18

LIVVOs are characterized respectively as $\sigma_{C^2-C^3}^*$ and $\pi_{C^2-C^3}^*$:



Note that in the case of the $\sigma_{C^2-C^3}^*$ LIVVO there are contributions from both s and p orbitals, while 98% of the $\pi_{C^2-C^3}^*$ LIVVO is comprised of p orbitals. The construction and analysis of the LIVVOs is fully automated and is performed only once after a ground state KS computation.

4.2.3 Analysis of OCDFT Particle Orbitals Using LIVVOs

The procedure used to generate the LIVVOs relies only on the information contained in the reference ground state determinant and is therefore applicable to any method. The focus of this work is to apply the LIVVO to OCDFT computations of K-edge spectra. First, we briefly summarize the features of OCDFT. In OCDFT,⁵³ a generalized Kohn–Sham picture is assumed, where to the n -th electronic state corresponds an auxiliary system of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho^{(n)}$. The wave function for the auxiliary system is a single Slater determinant, $|\Phi^{(n)}\rangle = |\phi_1^{(n)} \phi_2^{(n)} \cdots \phi_N^{(n)}\rangle$, where the set of orbitals $\{\phi_i^{(n)}\}$ are different for each electronic state. To avoid variational collapse, OCDFT enforces orthogonality conditions between all determinants:

$$\langle \Phi^{(m)} | \Phi^{(n)} \rangle = \delta_{mn} \quad \forall m, n \quad (4.10)$$

In this work we employ our constrained multiple hole/particle (CMHP) algorithm⁵⁴ for the solution of multiple orthogonally constrained excited states. The CMHP scheme fully accounts for relaxation of all orbitals, but rather than solving

for the exact minimum orthogonality conditions, it imposes sufficient orbital orthogonality conditions. Namely, we assume that the hole orbitals in core-excited states are localized and are orthogonal. It follows then that one-electron excited states can be characterized by fully relaxed hole $[\phi_h^{(n)}]$ and particle $[\phi_p^{(n)}]$ orbitals, which span the occupied and virtual spaces of the ground state wave function, respectively.

For every excited state, we analyze the character of the particle orbital $\phi_p^{(n)}$ by evaluating its overlap with each LIVVO (ψ_l^{LIVVO}):

$$\Omega_{pl}^{(n)} = |\langle \phi_p^{(n)} | \psi_l^{\text{LIVVO}} \rangle|^2. \quad (4.11)$$

The individual overlaps $\Omega_{pl}^{(n)}$ are used to assign the character of the particle orbital to the l -th LIVVO. In addition, we also define the total valence character $t_p^{\text{val},(n)}$ for any given particle orbital as the sum its overlap with all LIVVOs:

$$t_p^{\text{val},(n)} = \sum_{l=1}^{N_{\text{LIVVO}}} \Omega_{pl}^{(n)}. \quad (4.12)$$

Since the LIVVOs give an accurate representation of valence orbitals in the molecule, it is possible to quantify the most important local contributions to the particle orbital. While this offers a very robust description for localized valence particle orbitals, it is important to note that the particle orbital describing the excited state can also be fairly diffuse with respect to the molecular environment. Such an orbital will obviously have relatively low overlap with the LIVVOs described here. This can be a desirable feature of our method as it allows for immediate identification of spectral contributions that arise from more diffuse orbitals. However, any further classification of the diffuse nature or Rydberg character of virtual orbitals is outside the scope of the current study.

4.3 Computational Details

OCDFT is currently implemented as a plugin for the PSI4⁷⁵ *ab initio* quantum chemistry package. All core excitation energies and oscillator strengths are obtained using

the B3LYP^{76–80} functional and a triple zeta correlation-consistent basis set jun-cc-pVTZ,^{81,82} which is obtained by deleting diffuse functions from the highest subshell of every atom in the aug-cc-pVTZ basis set.⁸³ The two-electron integrals are approximated using density-fitting^{84–89} with the jun-cc-pVTZ-JKFIT auxiliary fitting basis.^{90,91} Basis set dependence of the LIVVO analysis is investigated using the cc-pVXZ and aug-cc-pVXZ (X=D,T,Q) family of basis sets.^{82,92} Intrinsic atomic orbitals were obtained using Robert Parrish’s implementation in PSI4.⁹³ The IAO scheme utilizes the MINAO basis set as used in Ref. 47, this set is simply a contracted subset of the cc-pVTZ basis. The geometries for ethanethiol, benzenethiol, water monomer, and water dimer were all optimized at the B3LYP/aug-cc-pVTZ level of theory.

Peak intensities for each transition from the ground state to excited state (n) are based on the oscillator strength (f_{osc}):

$$f_{osc} = \frac{2}{3} |\boldsymbol{\mu}_{n0}|^2 \omega_n \quad (4.13)$$

where ω_n is the excitation energy calculated in OCDFT and $\boldsymbol{\mu}_{n0}$ is the transition dipole moment vector. OCDFT does not provide a direct route toward computation of the transition dipole moments, however, these can be approximated using the noninteracting Kohn–Sham determinants:

$$\boldsymbol{\mu}_{n0} = \langle \Phi^{(n)} | \mathbf{r} | \Phi^{(0)} \rangle \quad (4.14)$$

where $\Phi^{(n)}$ is the KS determinant for the n -th excited state and \mathbf{r} is the position vector. Spectroscopic broadening effects are simulated by representing each transition using a Gaussian lineshape with a FWHM of 0.4 eV.

4.4 Results

4.4.1 Analysis of substituent effects in the spectra of thiols

Our first application of the LIVVO analysis for the classification of NEXAFS spectral features focuses on the effect of substituents on peak position and intensity. In

Table 4.2: Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of ethanethiol and benzenethiol. Experimental energies and assignments are taken from Ref. 57.

State	OCDFT			Experiment	
	Energy (eV)	f_{osc}	LIVVO Assignment	Energy (eV)	Assignment
Ethanethiol					
1	2472.9	0.0022	58% $\sigma_{\text{S}^3-\text{H}^9}^*$ 22% $\sigma_{\text{S}^3-\text{C}^2}^*$	2472.2	$\sigma_{\text{S}-\text{H}}^*$ Weak $\sigma_{\text{S}-\text{C}}^*$
2	2473.9	0.0011	57% $\sigma_{\text{S}^3-\text{C}^2}^*$ 18% $\sigma_{\text{S}^3-\text{H}^9}^*$	2473.1	$\sigma_{\text{S}-\text{C}}^*$
3	2475.6	0.0004	9% $\sigma_{\text{C}^2-\text{H}^8}^*$ 7% $\sigma_{\text{C}^2-\text{H}^7}^*$	2474.9	Poorly Defined
4	2476.2	0.0007	8% $\sigma_{\text{S}^3-\text{C}^2}^*$ 5% $\sigma_{\text{C}^2-\text{H}^8}^*$		
5	2476.1	0.0005	6% $\sigma_{\text{C}^1-\text{H}^6}^*$ 3% $\sigma_{\text{C}^2-\text{H}^7}^*$		
Benzenethiol					
1	2472.9	0.0023	61% $\sigma_{\text{S}^7-\text{H}^8}^*$ 22% $\sigma_{\text{S}^7-\text{C}^3}^*$	2472.4	$\sigma_{\text{S}-\text{H}}^*$ Weak $\sigma_{\text{S}-\text{C}}^*$
2	2474.2	0.0011	53% $\sigma_{\text{S}^7-\text{C}^3}^*$ 16% $\sigma_{\text{S}^7-\text{H}^8}^*$	2473.5	Weak $\pi_{\text{C}=\text{C}}^*$
3	2474.9	0.0001	68% $\pi_{\text{C}^{2,3,4}}^*$ 15% $\pi_{\text{C}^{4,5,6}}^*$	2475.3	$\sigma_{\text{S}-\text{C}}^*$
4	2475.3	0.0000	48% $\pi_{\text{C}^{1,2,6}}^*$ 34% $\pi_{\text{C}^{4,5,6}}^*$		
5	2476.0	0.0006	7% $\sigma_{\text{S}^7-\text{H}^8}^*$		

this section we compare the sulfur K-edge of ethanethiol ($\text{C}_2\text{H}_5\text{SH}$) and benzenethiol ($\text{C}_6\text{H}_5\text{SH}$). The OCDFT NEXAFS spectra for these two compounds are shown in Figure 4.2 along with the molecular structures and numbering scheme. Table 4.2 reports a comparison of experimental and theoretical excitation energies, OCDFT oscillator strengths, the LIVVO assignments, and total valence character for each state.

The gas-phase experimental excitation energies shown in Table 4.2 for ethanethiol represents the three peaks of the spectrum at 2472.17, 2473.11, and 2474.87 eV. Our calculated OCDFT spectrum is also characterized by three peaks with transition energies that are in good agreement with the experimental spectrum.⁵⁷ The first two

are a result of strong single transitions at 2472.9 and 2473.9 eV. The final peak is a combination of three weak transitions that occur at 2475.6, 2476.2, and 2476.1 eV.

The gas phase experimental spectrum of benzenethiol is characterized by three peaks at 2472.4, 2473.5, and 2475.3 eV. The theoretical spectrum has strong single transitions at 2472.9 eV and 2474.2 eV. The last peak is composed of weaker transitions at 2474.9, 2475.3, and 2476.0 eV. In total, this gives us errors that are on average less than 1.0 eV across these two molecules. This agreement with the experimental spectrum is consistent with that seen in previous studies with OCDFT^{54,95} and are obtained without application of shifts to match experimental values.

By analyzing the overlap of the particle orbitals with all orbitals in the LIVVO basis, it is possible to quantify the degree of localization along each bond and discern the origin of some features of the ethanethiol and benzenethiol spectra. The lowest energy transition in the NEXAFS spectrum of ethanethiol and benzenethiol occurs

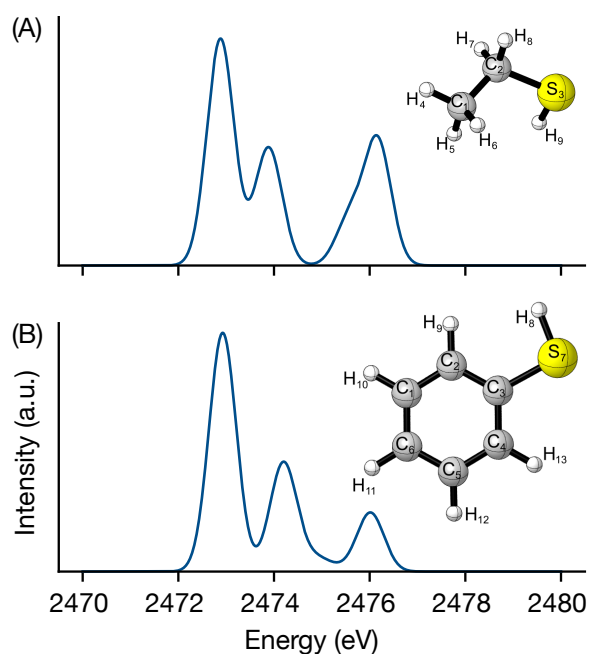


Figure 4.2: Sulfur 1s near-edge x-ray absorption spectra of thiols calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Spectra of (A) ethanethiol and (B) benzenethiol. Shown in the inset of each spectrum is the optimized structure for each molecule. Structures were plotted using CYLview software.⁹⁴

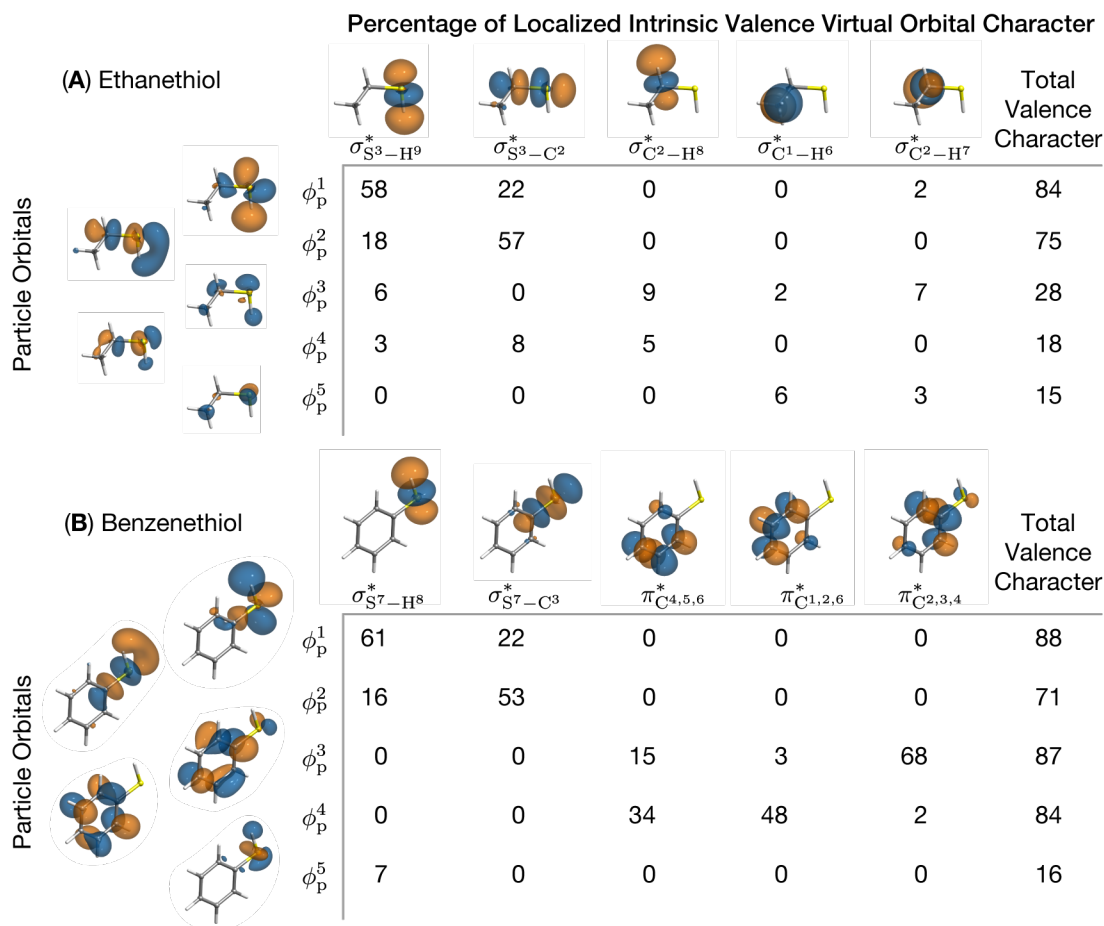


Figure 4.3: Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) ethanethiol (C_2H_5SH) and (B) benzenethiol (C_6H_5SH). Percentages shown represent the overlap of each particle orbital with the LIVVOs. Particle orbitals are numbered according to the calculated transitions reported in Table 4.2 while atom numberings correspond to those shown in Figure 4.2.

at at 2472.9 eV, and was attributed mostly to a $S\ 1s \rightarrow \sigma_{S-H}^*$ transition with a “weak” contribution from a $S\ 1s \rightarrow \sigma_{S-C}^*$ transition.⁵⁷ The LIVVO assignments, displayed in Figure 4.3, show that in both cases the first state (corresponding to the particle orbital ϕ_P^1) can be assigned to σ_{S-H}^* (50%) and σ_{S-C}^* (22%) character. Going from ethanethiol to benzenethiol there is a slight increase in the degree of localization along the thiol bond. This is evidenced by the σ_{S-H}^* character of the first excited state increasing from 58% in ethanethiol, to 61% in benzenethiol.

The second peak in the NEXAFS spectra of the thiols occurs at 2473.9 eV for

ethanethiol and at 2474.2 eV for benzenethiol. Again in this case, simple inspection of the particle orbitals (ϕ_p^2) shows orbital character along both the S–H and S–C bonds. The LIVVO assignments show that over 50% of the character of the particle orbital can be assigned to the σ_{S-C}^* orbital in each molecule, while less than 19% can be assigned to the σ_{S-H}^* orbital. Comparing our results with a previous study⁵⁶ that employed the Hartree–Fock static exchange (HF-STEX) method,^{96–101} our assignments agrees well for the case of ethanethiol. However, the OCDFT assignment is at odds with their assignment for benzenethiol where they attributed this peak feature to core transitions to π^* orbitals associated with the phenol ring. The discrepancy in this assignment likely stems from the differences in the energy ordering of the virtual orbitals between HF and DFT.

The third peak in the spectrum for both thiols is a combination of three low-intensity transitions. Of these, the most intense transition is predicted by OCDFT to be at 2476.2 eV for ethanethiol and 2476.0 eV for benzenethiol and corresponds to the particle orbital ϕ_p^4 . Considering the total valence characters shown in Figure 4.3 and the oscillator strengths shown in Table 4.2, we see a clear trend relating the valence character and the oscillator strength of the transition. In the case of ethanethiol, the low valence character of the transitions (28% or lower) suggests that the smaller oscillator strength of the third peak is due to the diffuse mixed Rydberg character of particle orbitals ϕ_p^3 , ϕ_p^4 , and ϕ_p^5 .

In the case of benzenethiol, the valence character of states ϕ_p^3 and ϕ_p^4 is high (87% and 84%, respectively) but these orbitals are mostly localized on the phenyl ring and give small contributions to the transition dipole moment. State ϕ_p^3 is a mixture of $\pi_{C^{2,3,4}}^*$ and $\pi_{C^{4,5,6}}^*$ at 68% and 15% respectively, while state ϕ_p^4 is a mixture of $\pi_{C^{1,2,6}}^*$ and $\pi_{C^{4,5,6}}^*$ at 48% and 34% respectively. This analysis suggests that the small overlap of these states with the sulfur atom is the cause of the small intensity of the corresponding transitions. State ϕ_p^5 of benzenethiol has a much lower total valence

character (16%), however, since it has 7% overlap with the $\sigma_{\text{S-H}}^*$ thiol bond orbital, the corresponding transition still has a modest intensity. We should note that thanks to the LIVVO analysis we were able to quantify the weak local contributions to the states that contribute to the third peak. Previous studies⁵⁶ were unable to assess the orbital character of this set of transitions.

4.4.2 Signatures of hydrogen bonding in the NEXAFS spectrum of the water dimer

Due to the critical role it plays in many processes in nature, hydrogen bonding is one of the most important noncovalent interactions in chemistry.¹⁰²⁻¹⁰⁶ The water dimer (H_2O)₂ is one of the simplest yet most important examples of hydrogen bonding in chemistry and has garnered plenty of attention from both theory and experiment.¹⁰⁷⁻¹¹⁸ This simple dimer is the basic building block of the structures of liquid water and ice, which have both been studied extensively with NEXAFS spectroscopy to uncover their underlying hydrogen bond coordination.¹¹⁹⁻¹²¹ A host of computational methods have been applied to study the effects of hydrogen bonding on the NEXAFS spectrum of water, including transition potential DFT (TPDFT),^{61-63,122,123} coupled-cluster singles and doubles (CCSD),^{9,124,125} and complex polarization propagator DFT (CPP-DFT).¹²⁶ Here, we utilize our LIVVO analysis in order to quantify these effects on the NEXAFS spectrum.

Water monomer. A host of computational methods have been applied to study the effects of hydrogen bonding on the NEXAFS spectrum of water, including transition potential DFT (TPDFT),^{61-63,122,123} coupled-cluster singles and doubles (CCSD),^{9,124,125} and complex polarization propagator DFT (CPP-DFT).¹²⁶ Here, we utilize our LIVVO analysis in order to quantify these effects on the NEXAFS spectrum. Figure 4.4A shows the NEXAFS spectrum for water monomer and the lowest-energy configuration of the water dimer. To facilitate our analysis, we separate the dimer spectrum into contributions from the hole localized on the oxygen accepting the hy-

Table 4.3: Transition energies, oscillator strengths (f_{osc}), and LIVVO assignments for all states calculated in the NEXAFS spectrum of the water monomer, dimer-A, and dimer-D.

State	Energy (eV)	f_{osc}	LIVVO Assignment
Monomer			
1	533.7	0.00811	42% $\sigma_{\text{O}_1-\text{H}_2}^*$ 42% $\sigma_{\text{O}_1-\text{H}_3}^*$
2	535.4	0.01839	38% $\sigma_{\text{O}_1-\text{H}_2}^*$ 38% $\sigma_{\text{O}_1-\text{H}_3}^*$
3	537.3	0.00786	Rydberg/Diffuse
4	537.8	0.00252	Rydberg/Diffuse
5	538.5	0.00167	6% $\sigma_{\text{O}_1-\text{H}_2}^*$ 6% $\sigma_{\text{O}_1-\text{H}_3}^*$
Dimer-A			
1	534.0	0.00737	41% $\sigma_{\text{O}_1-\text{H}_2}^*$ 41% $\sigma_{\text{O}_1-\text{H}_3}^*$
2	535.8	0.01862	39% $\sigma_{\text{O}_1-\text{H}_2}^*$ 39% $\sigma_{\text{O}_1-\text{H}_3}^*$
3	537.7	0.00552	14% $\sigma_{\text{O}_4-\text{H}_5}^*$ 9% $\sigma_{\text{O}_4-\text{H}_6}^*$
4	537.9	0.00529	4% $\sigma_{\text{O}_4-\text{H}_6}^*$ 3% $\sigma_{\text{O}_4-\text{H}_5}^*$
5	539.3	0.00282	12% $\sigma_{\text{O}_4-\text{H}_6}^*$ 8% $\sigma_{\text{O}_4-\text{H}_5}^*$
Dimer-D			
1	533.5	0.00605	52% $\sigma_{\text{O}_4-\text{H}_6}^*$ 16% $\sigma_{\text{O}_4-\text{H}_5}^*$
2	534.9	0.00878	24% $\sigma_{\text{O}_4-\text{H}_6}^*$ 18% $\sigma_{\text{O}_1-\text{H}_2}^*$
3	536.1	0.00574	3% $\sigma_{\text{O}_1-\text{H}_2}^*$ 3% $\sigma_{\text{O}_1-\text{H}_3}^*$
4	536.6	0.00284	2% $\sigma_{\text{O}_4-\text{H}_5}^*$ 2% $\sigma_{\text{O}_1-\text{H}_3}^*$
5	537.3	0.00623	25% $\sigma_{\text{O}_4-\text{H}_5}^*$ 5% $\sigma_{\text{O}_4-\text{H}_6}^*$

drogen bond (O_1 , denoted as dimer-A) and a spectrum for the hole localized on the oxygen donating the hydrogen bond (O_4 , denoted as dimer-D). These partial contributions are shown in panels B and C of Figure 4.4, respectively. Table 4.3 shows all of the calculated excitation energies, oscillator strengths and final LIVVO assignments for each state.

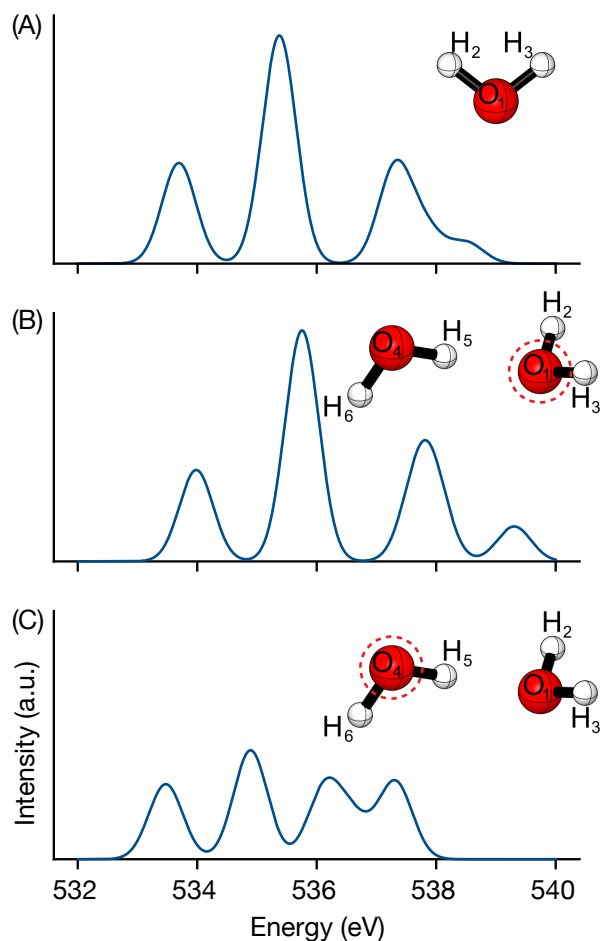


Figure 4.4: Oxygen 1s Near-edge x-ray absorption spectra of: (A) water monomer, (B) dimer-A, and (C) dimer-D calculated with OCDFT at the B3LYP/jun-cc-pVTZ level of theory. Shown in the inset of each spectrum is the optimized structure for each molecule. For the water dimer, we indicate the core-excited atom with a red circle. Structures were plotted using the CYLview software.⁹⁴

The OCDFT computed NEXAFS spectrum of water monomer at the O K-edge, shown in Figure 4.4, is defined by three characteristic peaks at 533.7, 535.4, and 537.3 eV respectively. The positions of these peaks are in excellent agreement with the gas-phase experimental peak features which occur at 534.0, 535.9, and 537.0 eV.¹²⁷ The first peak at 533.7 eV has a relative intensity of 0.44 and is characterized by the particle orbital ϕ_p^1 shown in Figure 4.5A, while the second peak at 535.4 is the most intense of the spectrum and is characterized by the particle orbital ϕ_p^2 . Two Rydberg transitions at 537.3 and 537.8 eV contribute to the third peak in the spectrum and

are accompanied by a higher energy shoulder feature at 538.5 eV.

The water monomer particle orbitals are shown in Figure 4.5A along with the overlap with the LIVVOs and their total valence character. The first two transitions in the water monomer are heavily localized resulting in total valence character of 85% and 75%, respectively. The three higher energy transitions are to more diffuse orbitals, as evidenced by the extremely low total valence percentages (in the range 0–12 %). Rydberg transitions that contain no overlap with the LIVVO basis due to their high degree of delocalization can not be assigned a well-defined valence character and are labeled “Rydberg/Diffuse” in Table 4.3. Analyzing the contributions of each LIVVO to the respective particle orbitals reveals that for this case, both the $\sigma_{\text{O}_1-\text{H}_2}^*$ and $\sigma_{\text{O}_1-\text{H}_3}^*$ LIVVOs contribute equally to each excited state. This is a characteristic feature of particle orbitals in the spectrum of the water monomer and will be useful in the comparison with the water dimer spectrum.

Water dimer, hydrogen bond acceptor. The spectra for localized excitations in the water dimer shown in panels B and C of Figure 4.4 allow us to discern features that arise from hydrogen bonding in the accepting and donating water molecules. Comparing panels A (monomer) and B (dimer-A) of Figure 4.4, the two spectra are very similar with three distinct peaks. The first two core excitations for the dimer-A spectrum occur at 534.0 and 535.8 eV and exhibit a 0.3 and 0.4 eV blueshift, respectively, when compared to the monomer. The oscillator strengths are also similar, with the second transition being the most intense in the spectrum, and the first transition having similar oscillator strength in the dimer-A (0.00737) and monomer (0.00811) spectra.

The similarities between these transitions becomes obvious when comparing the particle orbitals for the monomer and the dimer (shown in Figure 4.5). The particle orbitals for the first two transitions in the dimer-A spectrum have no contribution from the donating water molecule and are localized exclusively on the accepting water

molecule. In addition, the particle orbitals ϕ_p^1 and ϕ_p^2 for the dimer-A look nearly identical to those for the water monomer. These two particle orbitals are localized on the hydrogen atoms of the acceptor water and thus are weakly perturbed by the hydrogen bond. In contrast, the three higher energy transitions show some degree of delocalization onto the donating water molecule.

The LIVVOs can be used in an effort to quantify the degree of localization/delocalization in the particle orbitals involved. Water dimer contains four LIVVOs, two that are localized on the accepting water ($\sigma_{O_1-H_2}^*$ and $\sigma_{O_1-H_3}^*$), one that is localized on the free OH bond in the donor water ($\sigma_{O_4-H_6}^*$), and one that represents the hydrogen bonding interaction between the two water molecules ($\sigma_{O_4-H_5}^*$). While similar in orbital character to the other LIVVOs, $\sigma_{O_4-H_5}^*$ has contributions from the accepting water molecule which gives us a direct way to track the overlap of an excited state orbital with O-H antibonding interaction from the hydrogen bond.

The values reported in Figure 4.5B correspond to the percentage of overlap between each water dimer LIVVO and the particle orbitals contributing to the dimer-A spectrum. Considering that orbitals ϕ_p^1 and ϕ_p^2 are heavily localized on the accepting water, it is unsurprising that the first two states have equal contributions from the LIVVOs localized on the accepting water ($\sigma_{O_1-H_2}^*$ and $\sigma_{O_1-H_3}^*$). The total valence character of these states (84% and 77%) is similar to the total valence character of the first two particle orbitals in the monomer spectrum (85% and 75%), implying that the presence of the donor water has a negligible effect on ϕ_p^1 and ϕ_p^2 . It is encouraging to see that our LIVVO analysis parallels the uniformity seen in these states with regard to their orbital character, peak position, and peak intensity.

The three particle orbitals corresponding to higher energy transitions in the dimer-A spectrum (ϕ_p^3 , ϕ_p^4 , ϕ_p^5) have significant overlap with $\sigma_{O_4-H_5}^*$ and $\sigma_{O_4-H_6}^*$. It is in this higher energy region of the spectrum where we see the most significant distinctions between the dimer-A and monomer spectrum. The monomer spectrum shows a

shoulder feature at higher energy while in the dimer-A this shoulder feature is now a distinct peak at higher energy. Our analysis based on LIVVOs shows that states 3, 4, and 5 experience a stark increase in total valence character from 0%, 0%, and 12% respectively in the monomer to 23%, 7%, and 29% in the dimer-A. This drastic increase in valence character is caused by interactions with the donor water, in all cases the final LIVVO assignments (shown in Table 4.3) can be attributed to the two LIVVOs localized on the donor water. The spectral changes seen in this region of the dimer-A spectrum can be rationalized by the presence of distinct orbital character from the donor water. This result is particularly encouraging, since simple inspection of the particle orbitals give little indication of these contributions from the donor water. In fact, especially in the cases of ϕ_p^3 and ϕ_p^4 , these contributions look fairly insignificant. However our analysis reveals that these seemingly harmless contributions account for a majority, if not all of the total valence character for these states.

Water dimer, hydrogen bond donor. In contrast to dimer-A, the spectrum for dimer-D bears little resemblance to the monomer spectrum. Although there are still three distinct peak features, the oscillator strength of all transitions drop significantly. We also observe a redshift of all three peaks, the first peak exhibits a mild shift of 0.2 eV while the higher energy peaks show more dramatic shifts of 0.7 and 1.0 eV respectively. The stark differences in the dimer-D spectrum, can be rationalized using our particle orbital/LIVVO analysis. The particle orbitals for each state shown in Figure 4.5C clearly show significant differences from the monomer spectrum for the dimer-D. The particle orbital for the first state in the dimer-D spectrum has similar orbital character to the first state in the monomer spectrum. However, ϕ_p^1 for dimer-D is clearly perturbed by the formation of the hydrogen bond and localizes along the free OH bond. This phenomenon has been observed in previous studies,^{59,125} but the degree of localization has not been rigorously quantified.

The LIVVO overlaps reported in Figure 4.5C for ϕ_p^1 shows that 52% of the orbital

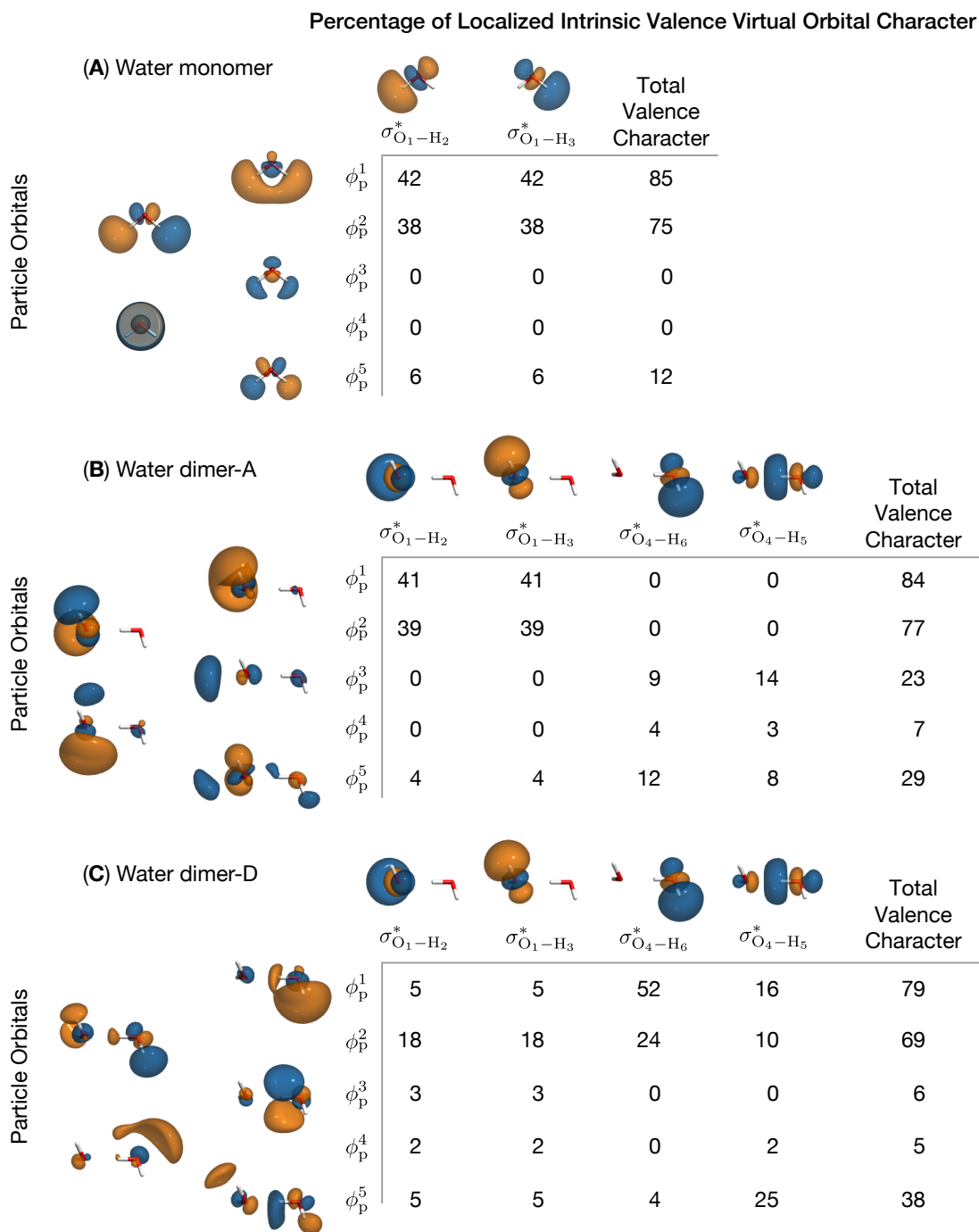


Figure 4.5: Particle orbitals, valence virtual orbitals, and total valence character for each of the five core transitions in the NEXAFS spectrum of (A) water monomer, (B) the water dimer at the oxygen accepting the hydrogen bond, and (C) the water dimer at the oxygen donating the hydrogen bond. Particle orbitals are numbered according to the calculated transitions reported in Table 4.3 while atom numberings correspond to those shown in Figure 4.4.

character lies along the $\sigma_{\text{O}_4-\text{H}_6}^*$ orbital with only 16% localized along the hydrogen bond and 10% of the orbital character coming from LIVVOs localized on the accepting water molecule. For ϕ_{p}^2 , a similar localization is observed, however, a significantly higher percentage (36%) of orbital character is accounted for by LIVVOs on the accepting water. This sharing of electron density with the accepting water molecule is indicative of the effect of hydrogen bonding on the dimer spectrum and has a direct correlation to the changes in intensity that we see in the simulated NEXAFS spectrum. Table 4.3 shows a drop in f_{osc} when comparing the monomer to dimer-D, with a modest decrease for state 1 (0.00811 \rightarrow 0.00605) and a far more drastic decrease for state 2 (0.01839 \rightarrow 0.00878). This intensity drop can be attributed to delocalization of the particle orbital and correlates well with the population of LIVVOs on the accepting water molecule (10% and 36%, respectively).

In conclusion, LIVVOs help to address the origin behind the stark differences in the dimer-A and dimer-D spectra. Perhaps the most revealing signature of hydrogen bonding in the water dimer is the decrease in intensity of the lowest two peaks in the dimer-D contribution. Past studies have reasoned that the differences in these two spectra arise from the polarization of the orbitals in water.^{59,125} Due to the large dipole moment the virtual orbital are expected to be polarized toward the two hydrogen atoms and the NEXAFS spectrum will show more pronounced changes for the water molecule donating a hydrogen bond than for the one accepting it. Our LIVVO analysis is consistent with this physical interpretation. In addition, it quantifies the degree of delocalization of the particle orbitals in terms of localized valence orbitals.

4.4.3 Basis Set Dependence of LIVVO Analysis

An important aspect to address concerning this method is its dependence upon the choice of basis set. To fully address this issue, there are three aspects of the method

that must be tested with respect to their basis set dependence: 1) How consistent are the LIVVO assignments, i.e. is there significant fluctuation in the LIVVO/IAO overlaps (eq 4.9)? 2) Are the LIVVOs assigned to the excited state consistent? and 3) How much fluctuation is there in the particle orbital/LIVVO overlaps (eq 4.11)? We will address these three points in this section, using the two LIVVOs used to assign the first two excited states in ethanethiol, $\sigma_{\text{S}^3-\text{H}^9}^*$ and $\sigma_{\text{S}^3-\text{C}^2}^*$.

First we will take a look at the assignment of the atomic character of the LIVVOs, Table 4.4 shows the overlap of each VVO with eight of the IAOs in ethanethiol. These values are extremely consistent regardless of the choice of basis set, only deviating by a maximum of 0.01.

Next we will evaluate if the LIVVO assignments of the particle orbital are consistent for different basis sets. For ethanethiol, ϕ_{p}^1 was assigned primarily to the $\sigma_{\text{S}^3-\text{H}^9}^*$ LIVVO with a weaker contribution from $\sigma_{\text{S}^3-\text{C}^2}^*$ while ϕ_{p}^2 received a complementary assignment. Regardless of the choice of basis set, the LIVVO classification of both excited states is consistent. In all cases, ϕ_{p}^1 has greater than 57% overlap with the $\sigma_{\text{S}^3-\text{H}^9}^*$ LIVVO and 25% or less overlap with the $\sigma_{\text{S}^3-\text{C}^2}^*$ LIVVO. Similarly, ϕ_{p}^2 has 57% or greater overlap with $\sigma_{\text{S}^3-\text{C}^2}^*$ and 25% or less overlap with $\sigma_{\text{S}^3-\text{H}^9}^*$. The characterization of the orbital character of excited states is by far the most important feature of this technique, the consistency seen in the data across different basis sets is very encouraging.

Lastly, we will address the consistency of the LIVVO/particle orbital overlap quantity presented in eq 4.11 with respect to choice of basis set. Table 4.4 shows that this value can potentially fluctuate $\pm 10\%$ with the choice of basis set. However, since we have shown that the formulation of the LIVVOs has a negligible dependence on the choice of basis set, we conclude that the basis set dependence seen in these values can solely be attributed to the basis set dependence of the particle orbitals. Although these values can change fairly significantly, it is encouraging that they do

Table 4.4: Basis set dependence of the IAO assignment of LIVVOs and overlap $[\Omega_{pl}^{(n)}]$ of the $\sigma_{S^3-H^9}^*$ and $\sigma_{S^3-C^2}^*$ LIVVOs with the first two particle orbitals of ethanethiol.

LIVVO	$\psi_7^{IAO}(C_s^2)$	$\psi_9^{IAO}(C_p^2)$	$\psi_{10}^{IAO}(C_p^2)$	$\psi_{13}^{IAO}(S_s^3)$	$\psi_{17}^{IAO}(S_p^3)$	$\psi_{18}^{IAO}(S_p^3)$	$\psi_{19}^{IAO}(S_p^3)$	$\psi_{25}^{IAO}(H_s^9)$	$\Omega_{pl}^{(1)}$	$\Omega_{pl}^{(2)}$
cc-pVDZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.08	0.23	0.00	0.15	0.54	63.7	22.3
$\sigma_{S^3-C^2}^*$	0.09	0.33	0.09	0.07	0.00	0.32	0.09	0.00	22.6	67.1
cc-pVTZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	59.3	23.9
$\sigma_{S^3-C^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	24.6	61.5
cc-pVQZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	57.8	22.8
$\sigma_{S^3-C^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	24.3	59.2
aug-cc-pVDZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.16	0.54	60.8	11.1
$\sigma_{S^3-C^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	15.3	63.1
aug-cc-pVTZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.08	0.22	0.00	0.15	0.54	57.9	14.7
$\sigma_{S^3-C^2}^*$	0.09	0.33	0.08	0.07	0.00	0.32	0.09	0.00	19.2	59.1
aug-cc-pVQZ										
$\sigma_{S^3-H^9}^*$	0.00	0.00	0.00	0.07	0.23	0.00	0.16	0.54	56.5	15.7
$\sigma_{S^3-C^2}^*$	0.08	0.33	0.09	0.07	0.00	0.32	0.09	0.00	20.4	57.3

not effect the overall interpretation of the excited state.

4.5 Conclusions

In this work, we present an automated method for the characterization of core-valence excited states. Our approach utilizes intrinsic atomic orbitals (IAOs) to derive a set of localized intrinsic valence virtual orbitals (LIVVOs) and assign their orbital character. LIVVOs are in turn used to classify particle orbitals calculated using orthogonality constrained density functional theory (OCDFT).

For molecular orbitals with dominant valence character, this analysis provides keen insight into the localized σ^* and π^* orbitals that contribute to it. For example, in the first two core excited states of ethanethiol, an electron is promoted to virtual orbitals that span both the S–C and S–H bonds. These contributions are difficult to discern by visual inspections of the MOs. For the first state, our analysis shows that the particle orbital is dominated by the $\sigma_{S^3-H^9}^*$ LIVVO (58%) with a smaller contribution

from the $\sigma_{\text{S}^3-\text{C}^2}^*$ (22%) LIVVO, revealing that this transition involves primarily the thiol bond. Instead, for the second state we find that weights are reversed, with the $\sigma_{\text{S}^3-\text{C}^2}^*$ LIVVO being the dominant contribution (57%).

We have also shown how LIVVOs may be useful to quantify differences in excited states due to changes in the molecular environment. For example, in the case of the water dimer we analyze the correlation between orbital localization due to hydrogen bonding and spectral features. The NEXAFS spectrum of water dimer at the oxygen K-edge is known to have distinct contributions from the oxygens at water donating (dimer-D) or accepting (dimer-A) the hydrogen bond. Analysis of the two lowest particle orbitals for excitations on dimer-A show equal localization along intramolecular OH bonds. However, the corresponding excitations for dimer-D are significantly different. In the lowest energy state, 70% of the electron is localized on the dimer-D, and this percentage decreases to 34% in the second state. These changes in degree of localization explain changes in intensity in the oxygen K-edge spectrum predicted by theory.

We have also considered the basis set dependence of the LIVVOs and the orbital analysis. Due to the robustness of the intrinsic atomic orbitals, the LIVVOs and our analysis of the orbital character are largely insensitive to the basis set size.

While our LIVVO analysis is applied here within the context of orthogonality constrained density functional theory, the method is general and can be used to decompose any set of virtual orbitals. Overall, our LIVVO-based analysis of particle orbitals is a first step toward creating a general, robust, and automatic method to assign the character of excited states.

Acknowledgments

This work was supported by start-up funds provided by Emory University. We would like to thank Dr. Robert M. Parrish for kindly providing us with the implementation of intrinsic atomic orbitals in Psi4. W.D.D. is supported by the National Science Foundation Graduate Research Fellowship under Grant No. 0000048655. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Bibliography

- [1] Westre, T. E.; Kennepohl, P.; DeWitt, J. G.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1997**, *119*, 6297–6314.
- [2] Löble, M. W. et al. *J. Am. Chem. Soc.* **2015**, *137*, 2506–2523.
- [3] Idrees, M.; Nadeem, M.; Sung, N.-E.; Asanova, T.; Shin, T. J. *Chem. Phys. Lett.* **2014**, *612*, 262–265.
- [4] Nelson, R. C.; Miller, J. T. *Catal. Sci. Technol.* **2012**, *2*, 461–470.
- [5] Stöhr, J. *NEXAFS Spectroscopy*; Springer Series in Surface Sciences; Springer, 1992.
- [6] Nelson, R. C.; Miller, J. T. *Catal. Sci. Technol.* **2012**, *2*, 461–470.
- [7] Milne, C. J.; Penfold, T. J.; Chergui, M. *Coord. Chem. Rev.* **2014**, *277*, 44–68.
- [8] Cabaret, D.; Bordage, A.; Juhin, A.; Arfaoui, M.; Gaudry, E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5619–5633.
- [9] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A.* **2012**, *85*, 022507.
- [10] Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 6735–6756.
- [11] Besley, N. A. *Chem. Phys. Lett.* **2012**, *542*, 42–46.
- [12] Peng, B.; Lestrangle, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.
- [13] Dutta, A. K.; Gupta, J.; Vaval, N.; Pal, S. *J. Chem. Theory Comput.* **2014**, *10*, 3656–3668.
- [14] Brabec, J.; Bhaskaran-Nair, K.; Govind, N.; Pittner, J.; Kowalski, K. *J. Chem. Phys.* **2012**, *137*, 171101.
- [15] Sen, S.; Shee, A.; Mukherjee, D. *Mol. Phys.* **2013**, *111*, 2625–2639.
- [16] Schirmer, J. *Phys. Rev. A.* **1982**, *26*, 2395.

- [17] Salpeter, E. E.; Bethe, H. A. *Phys. Rev.* **1951**, *84*, 1232–1242.
- [18] Barth, A.; Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. *Chem. Phys.* **1980**, *46*, 149–164.
- [19] Butscher, W.; Buenker, R. J.; Peyerimhoff, S. D. *Chem. Phys. Lett.* **1977**, *52*, 449–456.
- [20] Roemelt, M.; Neese, F. *J. Phys. Chem. A* **2013**, *117*, 3069–3083.
- [21] Asmuruf, F.; Besley, N. *Chem. Phys. Lett.* **2008**, *463*, 267–271.
- [22] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.
- [23] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.
- [24] Fernando, R. G.; Ballhoff, M. C.; Lopata, K. *J. Chem. Theory Comput.* **2015**, *11*, 646–654.
- [25] Rehr, J. J.; Albers, R. C. *Rev. Mod. Phys.* **2000**, *72*, 621–654.
- [26] Natoli, C. R.; Krüger, P.; Hatada, K.; Hayakawa, K.; Sébilleau, D.; Šipr, O. *J. Phys-Condens. Mat.* **2012**, *24*, 365501.
- [27] Joly, Y. *Phys. Rev. B* **2001**, *63*, 125120.
- [28] Slater, J. C. In *Statistical Exchange-Correlation in the Self-Consistent Field*; Löwdin, P.-O., Ed.; Adv. Quantum Chem.; Academic Press, 1972; Vol. 6; pp 1–92.
- [29] Stener, M.; Fronzoni, G.; Simone, M. d. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- [30] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039.
- [31] Lestrangle, P. J.; Nguyen, P. D.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.
- [32] Kuramoto, K.; Ehara, M.; Nakatsuji, H. *J. Chem. Phys.* **2005**, *122*, 014304.
- [33] Ohtsuka, Y.; Nakatsuji, H. *J. Chem. Phys.* **2006**, *124*, 054110.
- [34] Kolczewski, C.; Püttner, R.; Plashkevych, O.; Ågren, H.; Staemmler, V.; Martins, M.; Snell, G.; Schlachter, A. S.; Sant’Anna, M.; Kaindl, G.; Pettersson, L. G. M. *J. Chem. Phys.* **2001**, *115*, 6426–6437.
- [35] Wenzel, J.; Dreuw, A. *J. Chem. Theory Comput.* **2016**, *12*, 1314–1330.
- [36] Plasser, F.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2014**, *141*, 024106.
- [37] Plasser, F.; Bappler, S. A.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2014**, *141*, 024107.
- [38] Myhre, R. H.; Coriani, S.; Koch, H. *J. Chem. Theory Comput.* **2016**, *12*, 2633–2643.
- [39] Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M. *Chem. Phys.* **1982**, *71*, 51–64.
- [40] Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
- [41] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
- [42] Lee, M. S.; Head-Gordon, M. *Int. J. Quantum Chem.* **2000**, *76*, 169–184.
- [43] Subotnik, J. E.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2005**, *123*, 114108.

- [44] Auer, A. A.; Noojien, M. *J. Chem. Phys.* **2006**, *125*, 024104.
- [45] Laikov, D. N. *Int. J. Quantum Chem.* **2011**, *111*, 2851–2867.
- [46] Szczepanik, D.; Mrozek, J. *J Math Chem* **2013**, *51*, 2687–2698.
- [47] Knizia, G. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- [48] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.
- [49] Schmidt, M. W.; Hull, E. A.; Windus, T. L. *J. Phys. Chem. A* **2015**, *119*, 10408–10427.
- [50] West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. *J. Phys. Chem. A* **2015**, *119*, 10376–10389.
- [51] West, A. C.; Schmidt, M. W.; Gordon, M. S.; Ruedenberg, K. *J. Chem. Phys.* **2013**, *139*, 234107.
- [52] Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- [53] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- [54] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, 14360–14374.
- [55] Behyan, S.; Hu, Y.; Urquhart, S. G. *Chem. Phys. Lett.* **2014**, *592*, 69–74.
- [56] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2013**, *138*, 214302.
- [57] Behyan, S.; Hu, Y.; Urquhart, S. G. *J. Chem. Phys.* **2011**, *134*, 244304.
- [58] Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odellius, M.; Ogasawara, H.; Näslund, L. Å.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995–999.
- [59] Cavalleri, M.; Ogasawara, H.; Pettersson, L. G. M.; Nilsson, A. *Chem. Phys. Lett.* **2002**, *364*, 363–370.
- [60] Smith, J. D.; Cappa, C. D.; Messer, B. M.; Drisdell, W. S.; Cohen, R. C.; Saykally, R. J. *J. Phys. Chem. B* **2006**, *110*, 20038–20045.
- [61] Prendergast, D.; Galli, G. *Phys. Rev. Lett.* **2006**, *96*, 215502.
- [62] Iannuzzi, M. *J. Chem. Phys.* **2008**, *128*, 204506.
- [63] Leetmaa, M.; Ljungberg, M. P.; Lyubartsev, A.; Nilsson, A.; Pettersson, L. G. M. *J. Electron. Spectrosc. Relat. Phenom.* **2010**, *177*, 135–157.
- [64] Kühne, T. D.; Khaliullin, R. Z. *J. Am. Chem. Soc.* **2014**, *136*, 3395–3399.
- [65] Tse, J. S.; Shaw, D. M.; Klug, D. D.; Patchkovskii, S.; Vankó, G.; Monaco, G.; Krisch, M. *Phys. Rev. Lett.* **2008**, *100*.
- [66] Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *Science* **2004**, *306*, 851–853.
- [67] Smith, J. D.; Cappa, C. D.; Messer, B. M.; Cohen, R. C.; Saykally, R. J. *Science* **2005**, *308*, 793–793.

- [68] Chen, W.; Wu, X.; Car, R. *Phys. Rev. Lett.* **2010**, *105*.
- [69] Clark, G. N.; Cappa, C. D.; Smith, J. D.; Saykally, R. J.; Head-Gordon, T. *Mol. Phys.* **2010**, *108*, 1415–1433.
- [70] Vinson, J.; Kas, J. J.; Vila, F. D.; Rehr, J. J.; Shirley, E. L. *Phys. Rev. B* **2012**, *85*.
- [71] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.
- [72] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2638–2651.
- [73] Mulliken, R. *J. Chem. Phys.* **1955**, *23*, 1833–1840.
- [74] Mulliken, R. *J. Chem. Phys.* **1955**, *23*, 1841–1846.
- [75] Turney, J. M. et al. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 556–565.
- [76] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [77] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [78] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- [79] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [80] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [81] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- [82] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [83] Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.
- [84] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993–4999.
- [85] Schrader, D. M.; Prager, S. *J. Chem. Phys.* **1962**, *37*, 1456–1460.
- [86] Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys.* **1973**, *2*, 41 – 51.
- [87] Dunlap, B. I. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2113–2116.
- [88] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *Int. J. Quantum Chem., Symp.* **1977**, *12*, 81–87.
- [89] Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chemical Physics Letters* **1993**, *213*, 514–518.
- [90] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- [91] Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.
- [92] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- [93] Parrish, R. M.; Sherrill, C. D. *J. Am. Chem. Soc.* **2014**, *136*, 17386–17389.
- [94] Legault, C. Y. CYLview. 2009; <http://www.cylview.org>.

- [95] Verma, P.; Derricotte, W. D.; Evangelista, F. A. *J. Chem. Theory Comput.* **2016**, *12*, 144–156.
- [96] Otero, E.; Urquhart, S. G. *J. Phys. Chem. A* **2006**, *110*, 12121–12128.
- [97] Cooney, R. R.; Urquhart, S. G. *J. Phys. Chem. B* **2004**, *108*, 18185–18191.
- [98] Urquhart, S. G.; Turci, C. C.; Tyliszczak, T.; Brook, M. A.; Hitchcock, A. P. *Organometallics* **1997**, *16*, 2080–2088.
- [99] Urquhart, S. G.; Hitchcock, A. P.; Lehmann, J. F.; Denk, M. *Organometallics* **1998**, *17*, 2352–2360.
- [100] Urquhart, S. G.; Smith, A. P.; Ade, H. W.; Hitchcock, A. P.; Rightor, E. G.; Lidy, W. *J. Phys. Chem. B* **1999**, *103*, 4603–4610.
- [101] Kosugi, N.; Ueda, K.; Shimizu, Y.; Chiba, H.; Okunishi, M.; Ohmori, K.; Sato, Y.; Shigemasa, E. *Chem. Phys. Lett.* **1995**, *246*, 475–480.
- [102] Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294–300.
- [103] Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1–22.
- [104] Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117–4128.
- [105] Stiopkin, I. V.; Weeraman, C.; Pieniazek, P. A.; Shalhout, F. Y.; Skinner, J. L.; Benderskii, A. V. *Nature* **2011**, *474*, 192–195.
- [106] Pflugrath, J. W.; Quijcho, F. A. *Nature* **1985**, *314*, 257–260.
- [107] Feyereisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993–2997.
- [108] Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089–10094.
- [109] Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066–4073.
- [110] Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15–22.
- [111] Szalewicz, K.; Cole, S. J.; Kolos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662–3673.
- [112] Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765–1769.
- [113] Wang, Y.; Carter, S.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2008**, *128*, 071101.
- [114] Ch'ng, L. C.; Samanta, A. K.; Czakó, G.; Bowman, J. M.; Reisler, H. *J. Am. Chem. Soc.* **2012**, *134*, 15430–15435.
- [115] Shank, A.; Wang, Y.; Kaledin, A.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2009**, *130*, 144314.
- [116] Mukhopadhyay, A.; Cole, W. T. S.; Saykally, R. J. *Chem. Phys. Lett.* **2015**, *633*, 13–26.
- [117] Gómez, S.; Nafziger, J.; Restrepo, A.; Wasserman, A. *J. Chem. Phys.* **2017**, *146*, 074106.
- [118] Zhang, Z.; Jiang, W.; Wang, B.; Wang, Z. *Chemical Physics Letters* **2017**, *678*, 98–101.

- [119] Wilson, K. R.; Rude, B. S.; Catalano, T.; Schaller, R. D.; Tobin, J. G.; Co, D. T.; Saykally, R. J. *J. Phys. Chem. B* **2001**, *105*, 3346–3349.
- [120] Wilson, K. R.; Cavalleri, M.; Rude, B. S.; Schaller, R. D.; Nilsson, A.; Pettersson, L. G. M.; N Goldman,; Catalano, T.; Bozek, J. D.; Saykally, R. J. *J. Phys.: Condens. Matter* **2002**, *14*, L221.
- [121] Parent, P.; Laffon, C.; Mangeney, C.; Bournel, F.; Tronc, M. *J. Chem. Phys.* **2002**, *117*, 10842–10851.
- [122] Hetényi, B.; Angelis, F. D.; Giannozzi, P.; Car, R. *The Journal of Chemical Physics* **2004**, *120*, 8632–8637.
- [123] Cavalleri, M.; Odelius, M.; Nordlund, D.; Nilsson, A.; Pettersson, L. G. M. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2854–2858.
- [124] Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. *J. Chem. Theory Comput.* **2012**, *8*, 1616–1628.
- [125] Fransson, T.; Zhovtobriukh, I.; Coriani, S.; Wikfeldt, K. T.; Norman, P.; Pettersson, L. G. M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 566–583.
- [126] Ekström, U.; Norman, P. *Phys. Rev. A* **2006**, *74*.
- [127] Schirmer, J. *Phys. Rev. A* **1993**, *47*, 1136–1147.

Chapter 5

A Maximum Subspace Occupation Approach for the Study of the NEXAFS Spectra of Chemisorbed Organic Molecules Using Orthogonality Constrained Density Functional Theory: Pyrazine on Si(100) a Case Study

“We may, I believe, anticipate that the chemist of the future who is interested in structures with high molecular weight will come to rely upon a new structural chemistry, involving precise geometrical relationships among the atoms in the molecules”

Linus C. Pauling

Chapter Abstract

The near-edge X-ray absorption fine spectra (NEXAFS) at the carbon and nitrogen K-edge of pyrazine in the gas-phase and chemisorbed to the Si(100) surface are investigated using orthogonality constrained density functional theory (OCDFT). We introduce a novel approach to selectively target the 1s excitations of the adsorbate atoms based on atomic orbital subspace projection. This technique allows one to *a priori* specify an atomic orbital subset relevant to the organic adsorbate. Using

this approach we are able to simulate the NEXAFS spectrum of chemisorbed organic complexes at the K-edge of the adsorbate atoms. To analyze the calculated transitions we employ an orbital analysis based on localized intrinsic valence virtual orbitals (LIVVOs) in conjunction with OCDFT particle orbitals in order to provide a clear justification for every assignment. Utilizing the LIVVO analysis we are able to quantify the amount of π^*/σ^* mixing in each excited state and provide clear justification for changes in peak intensity.

5.1 Introduction

The adsorption of unsaturated hydrocarbons on the Si(100) surface has substantial technological significance:¹⁻³ exposed surface silicon dimers react with organic molecules forming covalent Si-C bonds which alters the reactivity of the silicon surface and opens up the possibility of fabricating novel semiconductor based molecular devices.⁴⁻⁷ This promise has inspired considerable investigation into the fundamental mechanisms that drive the chemisorption process.⁸⁻¹³ For this purpose, aromatic systems compose a particularly interesting class of molecules since their interaction with the surface often involves a myriad of unique bonding geometries that depend upon intrinsic properties such as the electronic and geometric structure of the molecule¹⁴⁻¹⁶ as well as extrinsic factors like temperature, coverage, and pressure.¹⁷⁻¹⁹ Aromatic molecules that contain a heteroatom can cause notable modifications the adsorption process. For example nitrogen-containing molecules have additional configurations

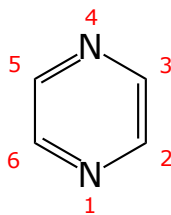


Figure 5.1: Molecular structure of pyrazine ($C_4H_4N_2$) with conventional numbering scheme.

that involve the donation of the nitrogen lone-pair electrons to the silicon.^{20–25} This occurs in the case of the pyrazine ($C_4H_4N_2$) molecule, a six membered aromatic heterocycle with electronic properties similar to those of benzene and pyridine. The ring system contains two nitrogen atoms at opposite ends (*para*), which increases the number of possible binding configurations upon molecular adsorption.¹⁵ Therefore the adsorption structure of pyrazine on the Si(100) surface is a very intriguing problem to study and has been the focus of many theoretical and experimental investigations.^{15,26–32}

High-resolution electron energy loss spectroscopy (HREELS) studies by Huang et al.²⁷ show that the adsorption of pyrazine on Si(100) occurs in a highly selective manner, directly bonding to the surface through the two *para* nitrogen atoms to form a 1,4-N-N-dihydropyrazine (DHP)-like structure (Figure 5.2b) attached to a single Si dimer. Density functional theory (DFT) cluster model calculations¹⁵ make a case for the N-end-on adsorbed pyrazine (Figure 5.2a) at low temperatures, and a species that is di- σ -bonded through the 2 and 5 carbons at elevated temperatures. DFT periodic slab studies by Jung and Kang²⁶ reveal that an isolated pyrazine molecule can bind

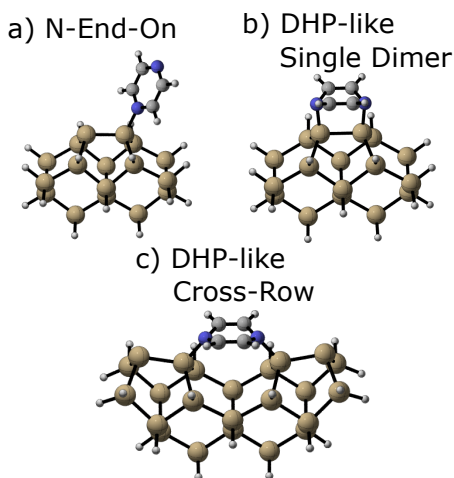


Figure 5.2: Schematic of three unique absorption modes of pyrazine ($C_4H_4N_2$) on Si(100), a) one with a single N atom attached at the Si surface dimer, and two structures that form a 1,4-N-N-dihydropyrazine (DHP)-like structure through forming two covalent Si-N bonds on a b) single dimer or c) across two dimer rows.

either in the N-end-on configuration or di- σ -bonded cross-row configuration (Figure 5.2c) with equal preference. At high coverages, the cross-row structure dominates and can form a linear molecular chain across the silicon dimer rows. This suggestion of the formation of a one-dimensional (1D) molecular chain is also in accordance with room temperature scanning tunneling microscopy (STM) and photoelectron diffraction (PED) studies by Shimomura et al.³¹ that suggest pyrazine forms this 1D molecular chain through a quasi-polymerization reaction through the Si dangling bonds. Pyrazine is one of the earliest examples of this type of self-assembled 1D molecular chain structure on a clean Si(100) surface.³⁰ This molecular assembly is particularly exciting in the case of pyrazine as it creates an ordered chain of planar C=C double bonds that could assist with further reaction at the semiconductor surface.

Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy is a powerful experimental technique to elucidate modifications induced on the electronic and geometric structure of a molecule upon adsorption to a surface.³³⁻³⁶ One of the powerful features of NEXAFS in application to this class of molecules lies in the angular dependence of the spectral features on the light polarization. By obtaining angle resolved spectra, one can glean useful information about the orientation of the molecule with respect to the surface.³⁷⁻³⁹ The fully-polarization resolved C K-edge spectrum of pyrazine on the surface of Si(100) was obtained by Han-Koo et al.²⁸ at 300K and 2L coverage.[?] Their analysis showed an angular dependence of the intensity of the C=C π^* resonance, resulting in an average tilt angle of the adsorbate with respect to the surface of approximately $34 \pm 5^\circ$. This result coupled with data from X-ray Photoelectron Spectroscopy, effectively eliminates all possible adsorption configurations except for the DHP-like cross-row configuration (See Figure 5.2c).

Analysis of the NEXAFS spectral features benefit greatly from a theoretical treatment. Calculations can provide useful insight into the relationship between the spectral features and structure.⁴¹ In this regard, NEXAFS simulations of adsorbed species

are particularly challenging because they require an accurate description of the adsorbate/substrate geometry and the core excited states of large systems.^{20,42,43} Finite cluster models are often employed to represent the adsorbed system as they are particularly adept at modeling phenomena that are localized on the surface or within the bulk of a solid.^{42,44} Interaction between the adsorbate and substrate model causes overlap of their molecular orbitals (MOs) and induces a rehybridization of the valence level. The position and intensity of features in the core spectrum will show modifications relative to the free molecule and can thus capture the influence that the surface has on the local electronic structure of the adsorbing molecule. Prior NEXAFS simulations for benzene and pyridine on Si(100) have employed finite models with as few as 9 to as many as 21 Si atoms.⁴⁵ The goal is to obtain an accurate description of the limited region around the adsorbate to capture the dominant effects on the NEXAFS spectrum due to the localization of core excited states near the core hole. Previous work by Rangan et al. has investigated the effect of increasing cluster size on the the core excitation spectrum of organic heterocycles adsorbed to Si(100).^{46,47} These studies show that increasing the number of Si atoms in the cluster model had a very small effect on the resulting spectrum with regard to peak position and intensity. Recent work by Romeo et al. on pyridine/Si(100) compared the results of independently optimized small clusters to larger clusters cut from periodic slab calculations.²⁰ These results highlighted differences in the individual polarized components of the spectrum, however, the effect on the total spectrum appears to be minimal. The computational efficiency and previous successful applications of finite cluster models has motivated their use for modeling the surface geometry in the current study with the understanding that any potential long-range interactions of the adsorbate with the extended surface will be neglected.

Aside from the treatment of the adsorbate/substrate geometry, computational approaches must calculate core excitation energies and properties in order to sim-

ulate the NEXAFS spectrum. The accuracy of calculated energies depends upon a reliable treatment of correlation, relaxation, and relativistic effects. Highly accurate many-body approaches have had considerable success in this area including coupled cluster theory,^{48–52} configuration interaction,^{53–55} and Green’s-function approaches.^{56,56–58} However, the computational cost associated with these methods limits their application to smaller systems. Considering the number of atoms necessary to construct cluster model surfaces it is desirable to consider methods that are computationally feasible for application to larger problems. Approaches that offer reduced computational cost include linear response time-dependent density functional theory (TDDFT),^{42,59,60} Hartree–Fock static exchange method, configuration interaction singles,⁶¹ Δ SCF, and real-time TD-DFT.⁶² We have recently proposed an alternative DFT based approach known as orthogonality constrained DFT (OCDFT)⁶³ and have shown that it yields highly accurate core excitation energies.^{64,65} It is a Δ SCF based Kohn–Sham DFT method which avoids variational collapse of the SCF solution by explicitly invoking an orthogonality constraint between the ground and excited states. Preliminary benchmark studies showed that OCDFT can accurately treat core excitations, providing absolute excitation energies that are more accurate than TDDFT.

Core excitation energies are calculated in OCDFT through the use of the constrained multiple hole particle (CMHP) algorithm. Similar to other excited state algorithms, CMHP is built to sequentially target excited state solutions starting from the lowest (highest) energy solution in a bottom-up (top-down) fashion. This presents an issue when targeting core excitations of organic adsorbates since the highest energy solutions are valence excitations and the lowest energy solutions are core excited states related to atoms of the cluster model. Thus, starting from either extrema requires the calculation of multiple unwanted solutions before reaching the adsorbate core. Having the ability to bypass these unwanted states and specifically target the core orbitals of the adsorbate is highly desirable.

Strategies to target specific core orbitals of interest are routinely employed to circumvent algorithmic challenges that exist in the calculation of core excited states. These methods typically fall into the category of either restricted excitation window (REW)^{62,66,67} or energy-specific (ES) techniques.^{60,68} These two methods differ by the way they solve the linear response equations. In the case of REW, the molecular orbital (MO) space from which excitations are allowed is restricted using either an orbital energy cutoff or MO Mulliken populations. In ES methods, the linear response equations are solved in the full MO space targeting eigenvalues that lie above a predefined energy threshold. To the best of our knowledge, similar strategies for targeting specific excitations have not been explored within the context of variational DFT methods.

Another theoretical challenge that is encountered by all methods is appropriately assigning transitions in chemisorbed complexes. Often times when classifying the NEXAFS spectra of gas-phase molecules it is sufficient to simply classify the transition as σ or π , which can easily be done by inspecting the character of the molecular orbitals involved in the final state. In the case of chemisorbed molecules it is also important to specify which atoms are major contributors to the final state orbital. This is crucial, since a large part of interpreting chemisorbed spectra is deciphering which peak features are a result of interactions with the surface and which features are largely localized on the adsorbate. Determining this through simple inspection of the MOs can be ambiguous, and thus a more quantitative approach toward determining this is desirable.

In this paper we introduce a new approach for targeting core-excitations in adsorbed molecules within OCDFT based on an atomic orbital subspace occupation analysis. Using this method we can circumvent the algorithmic difficulty of calculating core excitations from multiple low lying surface atoms and specifically focus on the adsorbate atoms of interest. By developing this method around specification

of atomic orbitals it requires little prior knowledge of the electronic structure of the molecule unlike other reduced subspace methods where exact energy ranges are required. We also address the issue of assigning transitions using a classification method based on localized intrinsic valence virtual orbitals.⁶⁹ We utilize these tools in order to analyze the NEXAFS spectrum of pyrazine chemisorbed on a Si(100) surface.

5.2 Method

Orthogonality Constrained Density Functional Theory

OCDFT is a time-independent variational density functional method⁷⁰ used to calculate electronic excited states. The original formulation of the theory can be found in ref [63](#), while details on its extension to treat core excitations and the algorithm used to calculate multiple excited states to simulate full NEXAFS spectra can be found in ref [64](#). In OCDFT, a generalized Kohn–Sham picture is assumed, where to the n^{th} electronic state there exists an auxiliary system of noninteracting electrons with wave function $\Phi^{(n)}$ and density $\rho^{(n)}$ that corresponds to the real interacting system $\Psi^{(n)}$. The wave function for the auxiliary system is a single Slater determinant, $|\Phi^{(n)}\rangle = |\phi_1^{(n)}\phi_2^{(n)}\cdots\phi_N^{(n)}\rangle$, where the set of orbitals $\{\phi_i^{(n)}\}$ are different for each electronic state.

$$\langle\Phi^{(m)}|\Phi^{(n)}\rangle = \delta_{mn} \quad \forall m, n \quad (5.1)$$

It can be shown, that given this constraint, one-electron excited states can be characterized by variationally optimized hole (ϕ_{h}) and particle (ϕ_{p}) orbitals, which must span the occupied and virtual spaces of the ground state wave function, respectively. In this work we employ our constrained multiple hole/particle (CMHP) algorithm for the solution of multiple orthogonally constrained excited states.⁶⁴ This allows us to readily compute multiple excited states by enforcing mutual orthogonality between the hole and particle orbitals of each subsequent excited state. At the same time, the

CMHP scheme fully accounts for relaxation of all orbitals.

Maximum Subspace Occupation

Here we introduce our approach for selectively targeting the core excitations of interest by combining OCDFT with a scheme to target specific hole orbitals. We start by considering a set of atomic orbital (AO) basis functions φ_s that are ordered by atom center, principal quantum number, and angular momentum. Based on this criteria, we can specify a subset of AO basis functions s that are of interest in a given system. For NEXAFS K-edge applications this is going to be the 1s core orbital(s) centered on the atom(s) of interest. The specified basis functions can be used to build an operator ($\hat{\Gamma}_s$) that projects onto the AO subset.

$$\hat{\Gamma}_s = \sum_s |\varphi_s\rangle\langle\varphi_s| \quad (5.2)$$

Utilizing this operator, we can evaluate the atomic orbital occupation number (Ω_i) of each molecular orbital ϕ_i within the desired subspace.

$$\Omega_i = \langle\phi_i|\hat{\Gamma}_s|\phi_i\rangle = \sum_s \langle\phi_i|\varphi_s\rangle\langle\varphi_s|\phi_i\rangle. \quad (5.3)$$

This is a number between 0 (no occupation in the AO subset) and 1 (full occupation in the AO subset). We have implemented this maximum subspace occupation (MSO) scheme in OCDFT for the selection of the hole orbital. Now instead of being chosen from the full occupied set, the hole orbitals ϕ_h are chosen from a subset of occupied orbitals ϕ_i with $\Omega_i \geq \omega$ where ω is a user-defined occupation threshold parameter. Figure 5.3 shows an example application of this subspace projection scheme to the calculation of carbon core excitations in carbon monoxide. In this example, the $C_{1s} \rightarrow \pi^*$ transition is targeted without needing to calculate any O_{1s} excitations beforehand. This projection scheme which is based on the nature of the atomic orbital provides a distinct advantage with respect to traditional REW and ES methods, as it requires

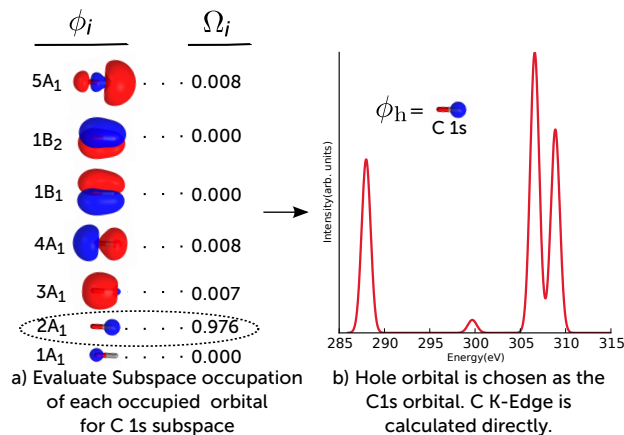


Figure 5.3: Example application of the maximum subspace occupation method within OCDFT to target the C 1s core excitations in CO at the B3LYP/3-21G level of theory: a) the 7 occupied orbitals of CO in C_{2V} symmetry, with the corresponding subspace occupation number (Ω_i) for a subspace composed of the C 1s atomic orbital. b) These projections are used to build the subset of occupied orbitals, in this example a hole threshold of 0.2 is used, and thus orbital $2A_1$ is chosen as the target hole orbital. This allows for direct access to the relevant C K-edge transitions.

no prior knowledge of the orbital energies or relevant energy range of absorption for a given system.

Table 5.1 shows the first 3 carbon core excitations of CCl_4 calculated with and without the maximum subspace occupation method. Comparison of these two approaches shows a negligible difference between the excitation energies (less than 0.001 eV) and oscillator strengths. The case of targeting the C K-edge of CCl_4 is comparable in this sense to targeting the C K-edge of chemisorbed organics. In order to compute three transitions of the C K-edge using the standard algorithm, one must first compute twelve excited states related to the Cl K-edge (three for each Cl atom). However, when applying the subspace projection technique, the C 1s orbitals can be targeted in an efficient manner without the need to calculate any Cl core excitations.

Table 5.1: First 3 carbon core excitation energies and oscillator strengths (f_{osc}) in a CCl_4 molecule calculated with and without the maximum subspace occupation (MSO) method in OCDFT. All calculations were performed using the B3LYP functional and aug-cc-pvdz basis set.

State	OCDFT		MSO-OCDFT	
	ω (eV)	f_{osc}	ω (eV)	f_{osc}
1	292.0002	0.0000	292.0005	0.0000
2	292.7093	0.0384	292.7097	0.0384
3	292.6669	0.0385	292.6659	0.0385

5.2.1 Assigning Transitions Based on Localized Intrinsic Valence Virtual Orbitals

In this section, we summarize our method for classifying excited states based on a novel localized orbital representation known as the *localized intrinsic valence virtual orbitals* (LIVVOs). Only a brief summary is given here, for full details on this technique we refer the reader to Ref. 69. This unique set of orbitals is a special case of the valence virtual orbitals (VVOs) developed by Rudenberg and coworkers.⁷¹ The goal of VVOs is to obtain chemically meaningful orbitals through evaluation of the valence character of the canonical molecular orbitals. This is done by identifying an accurate atomic minimal basis set (AAMBS) and evaluating the overlap of the canonical virtual orbitals with the AAMBS functions $\psi_{\rho}^{\text{AAMBS}}$ such that:

$$(\mathbf{S})_{a\rho} = \langle \phi_a | \psi_{\rho}^{\text{AAMBS}} \rangle. \quad (5.4)$$

For our implementation we have chosen to use the intrinsic atomic orbitals of Kinizia⁷² as the necessary AAMBS functions. Following the suggestion of Ref. 73 a singular value decomposition (SVD) is performed on \mathbf{S} :

$$\mathbf{S} = \mathbf{U}\sigma\mathbf{V}^{\dagger} \quad (5.5)$$

to yield the orthogonal transformation matrices \mathbf{U} and \mathbf{V} . These are rotations of the canonical virtual space and the AAMBS space, respectively that bring the two orbital sets into maximum coincidence. The transformation matrix \mathbf{U} is then applied to the

virtual set in the following way:

$$|\psi_\nu^{\text{VVO}}\rangle = \sum_a^{N_{\text{vir}}} |\phi_a\rangle U_{a\nu}, \nu = 1, 2, \dots, N_{\text{VVO}} \quad (5.6)$$

where N_{vir} is the number of virtual orbitals while N_{VVO} is the number of VVOs. Finally, the set of LIVVOs (ψ_l^{LIVVO}) are formed by using Pipek-Mezey localization,⁷⁴ and we arrive at a set of localized VVOs that span the antibonding interactions of the molecular environment.

For every excited state in OCDFT, we analyze the character of the particle orbital $\phi_p^{(n)}$ by evaluating its overlap with each LIVVO:

$$\Omega_{\text{pl}}^{(n)} = |\langle \phi_p^{(n)} | \psi_l^{\text{LIVVO}} \rangle|^2 \quad (5.7)$$

The individual overlaps $\Omega_{\text{pl}}^{(n)}$ are used to assign the character of the particle orbital to the l -th LIVVO. In addition, we also define the total valence character $t_p^{\text{val},(n)}$ for any given particle orbital as the sum of its overlap with all LIVVOs:

$$t_p^{\text{val},(n)} = \sum_l^{N_{\text{LIVVO}}} \Omega_{\text{pl}}^{(n)}. \quad (5.8)$$

With this set of orbitals it is possible to quantify the most important local contributions to the particle orbitals to provide a specific description of its character to make a robust assignment for the excited state.

5.3 Computational Details

The gas-phase geometry of pyrazine was optimized using the B3LYP⁷⁵⁻⁷⁸ functional and the 6-31G*⁷⁹⁻⁸¹ basis set. All core excited states were calculated using OCDFT in the PSI4 *ab-initio* quantum chemistry package.⁸² Core excited state calculations on the free molecule were done at the PBE/cc-pVTZ^{83,84} level of theory. Pyrazine is then attached in the double dimer bridging configuration to a Si₂₃H₂₄ cluster model of the Si(100) surface reconstruction. Covalent Si-N bonds are formed between adjacent

dimer rows and the nitrogen atoms of pyrazine. Core excitation calculations on chemisorbed pyrazine utilize the PBE functional⁸³ and a custom basis set where the cc-pVTZ basis set⁸⁴ is used for all carbon and hydrogen atoms and the 6-31G basis set⁸⁵ for all silicon atoms. The AO subspace is chosen as the C 1s or N 1s orbitals to simulate their respective K-edges with a hole threshold of 0.2.

We note that pyrazine is a highly symmetrical molecule containing symmetry equivalent C and N atoms. Molecules that contain symmetry equivalent cores present a challenging theoretical issue.^{86,87} In practice, symmetry restricted calculations produce a core hole that is delocalized evenly amongst all symmetry equivalent atoms. In contrast, a symmetry unrestricted calculation may produce core holes that are localized on each individual atom. For calculations on the free pyrazine molecule, we have studied both the symmetry restricted and unrestricted solutions. To obtain a state where the core hole is localized on each atom we utilize a wavefunction with broken spatial and spin symmetry by mixing the coefficients of the alpha HOMO and LUMO orbitals of pyrazine.

For spectral simulations, the peak intensities are based on the oscillator strength (f_{osc}):

$$f_{osc} = \frac{2}{3} |\mu_{n0}|^2 \omega_n \quad (5.9)$$

where μ_{n0} represents the transition dipole moment for transition from the ground state to excited state n and ω_n is the excitation energy for that transition. The spectrum is simulated by centering Gaussian functions at each excitation energy with a peak height scaled by the oscillator strength of that transition. The peak is then broadened by a constant value in order to simulate natural spectroscopic broadening effects, for the purpose of pyrazine we utilized a Gaussian width of 0.3 eV.

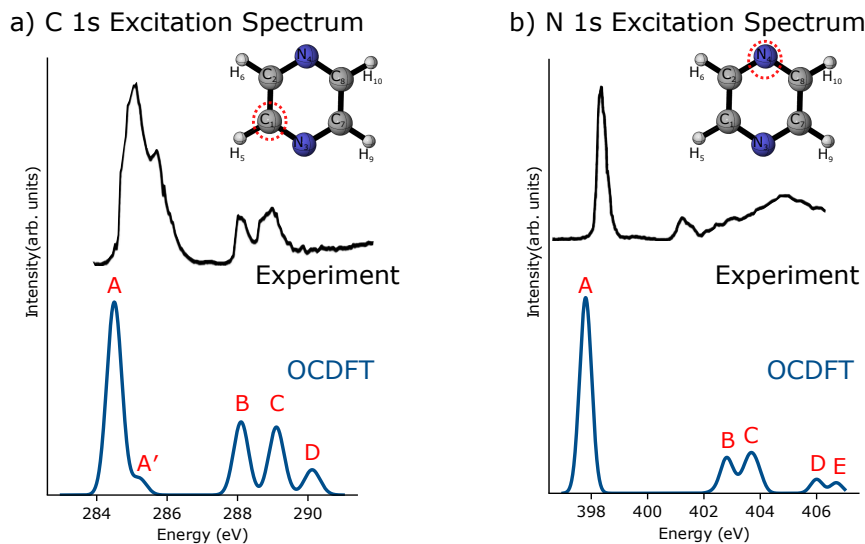


Figure 5.4: Near-edge X-ray absorption spectrum of pyrazine ($C_4H_4N_2$) at the a) C K-edge and b) N K-edge computed with OCDFT (blue line) at the PBE/cc-pVTZ level of theory. Experimental results (black line) are obtained from Ref. 88 and appear here in their original peak position with no applied energy shifts. The pyrazine geometry optimized at the B3LYP/6-31G* level is shown in the inset, the atoms circled (C_1 and N_4) are the core holes considered in further discussion of the results in this section.

5.4 Results and discussion

NEXAFS Spectrum of Gas-Phase Pyrazine

In this section we present an OCDFT simulation of the carbon and nitrogen 1s excitation spectrum for the free pyrazine molecule. Figure 5.4 shows the experimental and calculated NEXAFS spectrum along with corresponding peak labels. The relevant particle orbitals and the LIVVOs used to characterize each excited state are shown in Figure 5.5. Calculated excitation energies, oscillator strengths, and assignments are shown in Table 5.2.

The core excitation spectrum of pyrazine follows a pattern common to other unsaturated molecules: a dominant edge onset associated with a π^* resonance localized near the core hole, followed by a series of weaker transitions that are a mix of σ^* transitions, Rydberg transitions, and π^* transitions localized away from the core hole. The

Table 5.2: Calculated (PBE/cc-pVTZ) core excitation energies, oscillator strengths, and LIVVO assignments for the C and N 1s excitation spectrum of free pyrazine. The LIVVO assignments shown are the two with the highest percentage overlap with the particle orbital. For each calculated transition we report the π^*/σ^* mixing ratio (see text for details) and the spectral feature that this state contributes to (see Figure 5.4).

State	energy (eV)	osc	LIVVO Assignments	π^*/σ^* Mixing	t_p^{val}	Peak Contribution
C 1s Core Excitations						
1	284.5	0.0207	78.2% $\pi_{\text{C}_1-\text{N}_3}^*$ 19.3% $\pi_{\text{C}_7-\text{C}_8}^*$	100% π^*	97.9%	A
2	285.2	0.0020	51.1% $\pi_{\text{C}_1-\text{N}_3}^*$ 43.9% $\pi_{\text{C}_7-\text{C}_8}^*$	100% π^*	97.3%	A'
3	288.1	0.0083	45.4% $\sigma_{\text{C}_1-\text{H}_5}^*$ 12.0% $\sigma_{\text{C}_7-\text{H}_9}^*$	100% σ^*	64.1%	B
4	289.1	0.0042	28.6% $\pi_{\text{N}_3-\text{C}_7}^*$ 7.4% $\sigma_{\text{C}_1-\text{H}_5}^*$	65% $\pi^*/35\%\sigma^*$	76.4%	C
5	289.1	0.0041	26.0% $\pi_{\text{C}_7-\text{C}_8}^*$ 7.9% $\sigma_{\text{C}_1-\text{H}_5}^*$	61% $\pi^*/39\%\sigma^*$	74.6	C
6	289.4	0.0003	25.7% $\sigma_{\text{C}_2-\text{H}_6}^*$ 14.6% $\sigma_{\text{C}_7-\text{H}_9}^*$	100% σ^*	54.0	C
7	290.1	0.0027	20.5% $\sigma_{\text{C}_1-\text{N}_3}^*$ 12.8% $\sigma_{\text{C}_2-\text{N}_4}^*$	100% σ^*	59.5	D
8	290.2	0.0037	17.4% $\sigma_{\text{C}_8-\text{H}_{10}}^*$ 13.7% $\sigma_{\text{C}_7-\text{H}_9}^*$	100% σ^*	49.9	D
N 1s Core Excitations						
1	397.8	0.0211	59.4% $\pi_{\text{C}_2-\text{N}_4}^*$ 34.8% $\pi_{\text{C}_1-\text{N}_3}^*$	100% π^*	98.4%	A
2	399.3	0.0000	60.3% $\pi_{\text{C}_7-\text{N}_8}^*$ 22.2% $\pi_{\text{N}_4-\text{C}_8}^*$	100% π^*	96.9%	
3	402.4	0.0003	19.3% $\sigma_{\text{C}_2-\text{H}_6}^*$ 19.3% $\sigma_{\text{C}_8-\text{H}_{10}}^*$	100% σ^*	64.1%	B
4	402.8	0.0042	38.3% $\pi_{\text{C}_1-\text{N}_3}^*$ 6.0% $\sigma_{\text{C}_2-\text{H}_6}^*$	88% $\pi^*/12\%\sigma^*$	85.4	B
5	403.2	0.0010	17.9% $\sigma_{\text{C}_8-\text{H}_{10}}^*$ 8.9% $\pi_{\text{C}_1-\text{N}_3}^*$	29% $\pi^*/71\%\sigma^*$	58.7%	C
6	403.6	0.0034	21.6% $\sigma_{\text{C}_2-\text{N}_4}^*$ 21.6% $\sigma_{\text{N}_4-\text{C}_8}^*$	100% σ^*	68.3	C
7	403.8	0.0001	13.2% $\sigma_{\text{C}_7-\text{H}_9}^*$ 12.9% $\sigma_{\text{C}_1-\text{H}_5}^*$	100% σ^*	59.5	D
8	403.9	0.0024	16.0% $\sigma_{\text{C}_1-\text{H}_5}^*$ 15.6% $\sigma_{\text{C}_7-\text{H}_9}^*$	100% σ^*	52.4	D

experimental C 1s spectrum shown in Figure 5.4a follows this general trend. The experimental spectrum is characterized by four main peaks at 285.3, 285.8, 288.2, and 289.1 eV. The simulated spectrum using OCDFT agrees well with corresponding peak features at 284.5, 285.2, 288.1, and 289.1 eV respectively.

The OCDFT spectrum for the C K-edge was calculated considering excitations

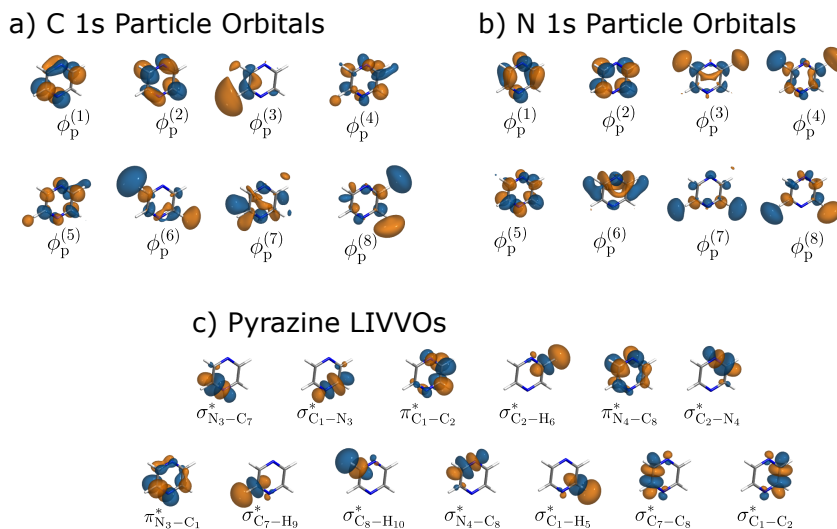


Figure 5.5: Particle orbitals for the core excited states that contribute to the a) carbon and b) nitrogen 1s excitation spectrum. Also shown are c) the LIVVOs that are used to analyze the particle orbitals.

from carbon atom C_1 (see inset in Figure 5.4). Due to the symmetry of the molecule, it is obvious that the equivalent carbon cores, would produce identical contributions to the core excitation spectrum. Limiting our discussion to a single core also aides in our discussion of localization when analyzing our LIVVO assignments. The first peak A in the carbon 1s spectrum occurs at 284.5 eV with a small shoulder feature A' occurring at 285.2 eV. These features can both be attributed to π^* transitions that are a combination of the $\pi_{C_1-N_3}^*$ and $\pi_{C_7-C_8}^*$ LIVVOs. In the case of Peak A, 78.2% of the particle orbital can be attributed to $\pi_{C_1-N_3}^*$ which is localized near the core hole on atom C_1 . On the contrary, the particle orbital for the shoulder feature A' , only 51.1% can be attributed to $\pi_{C_1-N_3}^*$, while 43.9% is attributed to $\pi_{C_7-C_8}^*$ which is localized away from the core hole. The calculated oscillator strength is sensitive to this change in localization, with state 1 having an oscillator strength of 0.0207 and state 2 having an oscillator strength ten times smaller (0.0020). State 3 is the sole contribution to Peak B, with an excitation energy of 288.1 and an orbital character that can be assigned as $\sigma_{C_1-H_5}^*$.

States 4 and 5 are quasi-degenerate both yielding excitation energies of 289.1 eV.

Looking at their particle orbitals $\phi_p^{(4)}$ and $\phi_p^{(5)}$ in Figure 5.5, it is clear that they have mixed π^* and σ^* character. This orbital mixing was noted, but not quantified, in a previous study.⁸⁸ The LIVVO analysis attributes 65% π^* and 35% σ^* character to state 4, and 61% π^* and 39% σ^* character to state 5. This similarity in the composition of their orbital character explains the similarities between their energy and oscillator strength. A weak transition ascribed to a $\sigma_{C_2-H_6}^*$ at 289.4 eV also contributes to Peak C. States 7 and 8 compose peak D at 290.1 and 290.2 eV. They are assigned to $\sigma_{C_1-N_3}^*$ and $\sigma_{C_8-H_{10}}^*$ respectively.

The N 1s core excitation spectrum shown in Figure 5.4b shows a similar general structure to the C 1s spectrum. Peak A occurs at 397.8 eV and can be assigned to $\pi_{C_2-N_4}^*$, which is the π^* orbital localized near the core hole. Interestingly, state 2 (399.3 eV) has null oscillator strength and does not contribute to the N-edge spectrum. The LIVVO analysis shows that 60.3% of this particle orbital can be attributed to the $\pi_{C_7-N_8}^*$ orbital, which is localized away from the nitrogen core hole at N₄. Also, the particle orbital $\phi_p^{(2)}$ shown in Figure 5.5b has no contribution localized at the nitrogen atoms. State 3 is a small contribution to Peak B at 402.4 eV attributed evenly to $\sigma_{C_2-H_6}^*$ and $\sigma_{C_8-H_{10}}^*$. As with State 2, much of the orbital character is localized away from the nitrogen resulting in a much weaker transition. As seen in the C 1s spectrum, States 4 and 5 are of mixed π^* and σ^* character. However, unlike the mixed states in the C spectrum, the composition of these orbitals in the N spectrum are unique. State 4, which contributes to Peak B at 402.4 eV, contains 88% π^* character and only 12% σ^* character. This state is assigned primarily to the $\pi_{C_1-N_3}^*$ orbital localized near the core hole, resulting in an appreciable oscillator strength of 0.0042. Although State 5 is also a mixed state, it contains far more σ^* character (29% π^* character and 71% σ^*). This state can be assigned to the $\sigma_{C_8-H_{10}}^*$ LIVVO, which explains its weak intensity in the N 1s spectrum. State 6 is a more intense contribution to Peak C at 403.6 eV and is assigned to $\sigma_{C_2-N_4}^*$ and $\sigma_{C_8-N_4}^*$, the two C-N σ -bonds associated with the core

hole N atom. States 7 and 8 contribute to Peak D at 403.8 and 403.9 eV respectively and are composed of transitions to σ^* orbitals associated with C-H bond along the ring system.

NEXAFS Spectrum of Chemisorbed Pyrazine on Si(100)

The OCDFT simulated NEXAFS spectrum of pyrazine chemisorbed on the Si(100) surface is shown in Figure 5.6. The LIVVOs shown in Figure 5.7 are only those that are directly involved with the adsorbed pyrazine molecule. In total, this system produces 70 LIVVOs. However many of these are $\sigma_{\text{Si-Si}}^*$ and $\sigma_{\text{Si-H}}^*$ orbitals related to the Si(100) cluster model. When comparing the LIVVOs of the chemisorbed complex to that of free pyrazine, many of the orbitals are remain unchanged except for three critical differences. Firstly, the most obvious difference is the appearance of the $\sigma_{\text{N}_3-\text{Si}_{11}}^*$ and $\sigma_{\text{N}_4-\text{Si}_{12}}^*$ LIVVOs to reflect the covalent bonds formed between the cluster and the nitrogens on pyrazine. Secondly, the π system of the molecule has changed

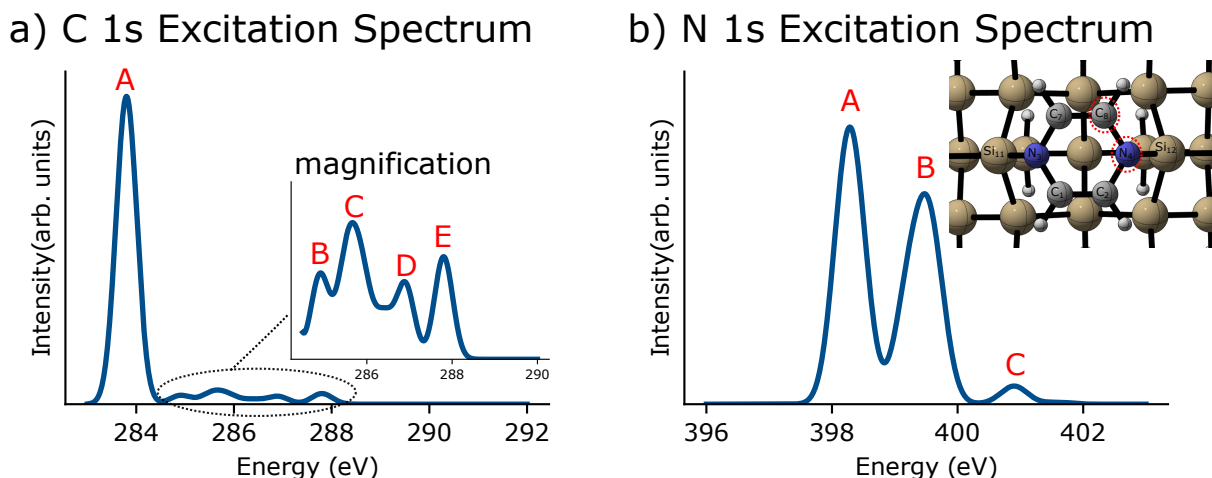


Figure 5.6: Near-edge X-ray absorption spectrum of pyrazine ($\text{C}_4\text{H}_4\text{N}_2$) chemisorbed on a $\text{Si}_{23}\text{H}_{24}$ cluster model of Si(100) at the a) C K-edge and b) N K-edge computed with OCDFT at the PBE level. All C, N, and H atoms use the cc-pVTZ basis set while the 6-31G basis set is used for the Si atoms of the cluster model. The inset in the C spectrum magnifies the σ^* manifold of the spectrum. The pyrazine geometry optimized at the B3LYP/6-31G* level is shown in the inset in the N spectrum, the atoms circled (C_8 and N_4) are the core holes considered in further discussion of the results in this section.

and the LIVVOs reflect that, the π^* orbitals are not localized along the N-C bonds, as is the case in free pyrazine, they are now localized along each C=C double bond with clear contributions from the N lone pair electrons. Lastly, the $\sigma_{\text{Si}_{13}-\text{Si}_{14}}^*$ represents the dangling bonds that exist at each open Si site on the dimers. These changes in the character of the LIVVOs gives us some insight into the effect of the cluster on the adsorbate.

Since pyrazine maintains two double bonds upon binding on the surface, the NEXAFS spectrum can be expected to maintain the same general spectral pattern of a sharp π^* peak and a weaker σ^* manifold. The experimental NEXAFS spectrum of chemisorbed pyrazine reports the initial spectral feature at 285.9 attributed to a π^* excitation. Our results show an initial excitation at 279.4 eV that is attributed to a $\sigma_{\text{Si}_{13}-\text{Si}_{14}}^*$, which is the orbital associated with the open Si atoms. The appearance of this state in our calculations is purely an artifact of the cluster model. In reality, pyrazine forms a 1D molecular chain along the surface, leaving few open Si sites on the surface and thus quenching this state. State 2 at 283.8 eV agrees well with the experimental onset feature, is assigned to π^* LIVVVOs, and is the sole contribution

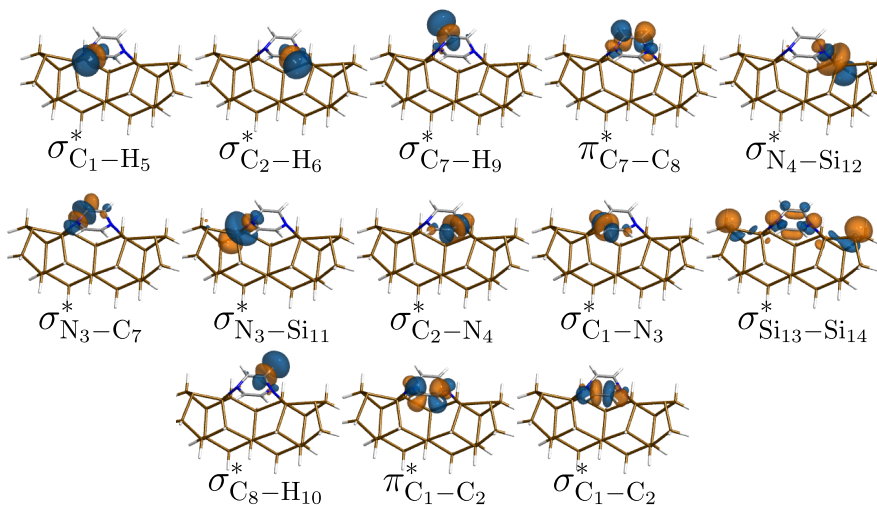


Figure 5.7: LIVVOs used to analyze particle orbitals in pyrazine ($\text{C}_4\text{H}_4\text{N}_2$) chemisorbed on a $\text{Si}_{23}\text{H}_{24}$ cluster model of Si(100).

to Peak A in the spectrum. Considering the core hole at atom C₈ State 2 is localized heavily toward the core hole, with 73.4% of the orbital character attributed to the $\pi_{C_7-C_8}^*$ LIVVO. This result is similar to the localization along a single π^* bond seen in the gas phase spectrum where the first state had 78.2% overlap with a single π^* LIVVO. Hence, the localization of this π^* state is not significantly affected upon adsorption to the surface. Due to this, there is a large difference in intensity at the C K-edge of chemisorbed pyrazine between the π^* peak and the σ^* manifold, that didn't exist in the spectrum of free pyrazine. The crucial difference here is the interaction and subsequent rehybridization of the pyrazine valence orbitals with the Si cluster. Looking at the particle orbital $\phi_p^{(2)}$ it is clear that this initial π^* state interacts very little with the Si cluster. Meanwhile every other particle orbital has significant contribution from Si cluster orbitals. This can also be seen quantitatively by looking at the LIVVO assignments in Table 5.3, for the π^* state (State 2), a single VVO constitutes 76% of the total valence character, in contrast no other state has a single VVO contribution of over 20%. As a result, the σ^* manifold for the C 1s spectrum is orders of magnitude less intense than the π^* feature.

The N 1s spectrum shown in Figure 5.6b, does not show the same stark contrast in intensity as the C spectrum. This is due to the fact that the π^* orbitals are now localized away from the N atoms upon the formation of the covalent Si-N bonds. In fact, in contrast to all other spectra considered here, the pure π^* state (State 10) is the highest in energy at 401.6 eV and has an oscillator strength of zero. Interestingly this is not the only state to provide no intensity to the spectrum. States 6 and 7 also provide no intensity to the spectrum due to their interaction with the cluster, both states have no contributions from LIVVOs that are localized on the pyrazine molecule and their respective particle orbitals in Figure 5.8b ($\phi_p^{(6)}$ and $\phi_p^{(7)}$) clearly shows no interactions from pyrazine.

Peak A of the N 1s spectrum has contributions from States 1 and 2, and corre-

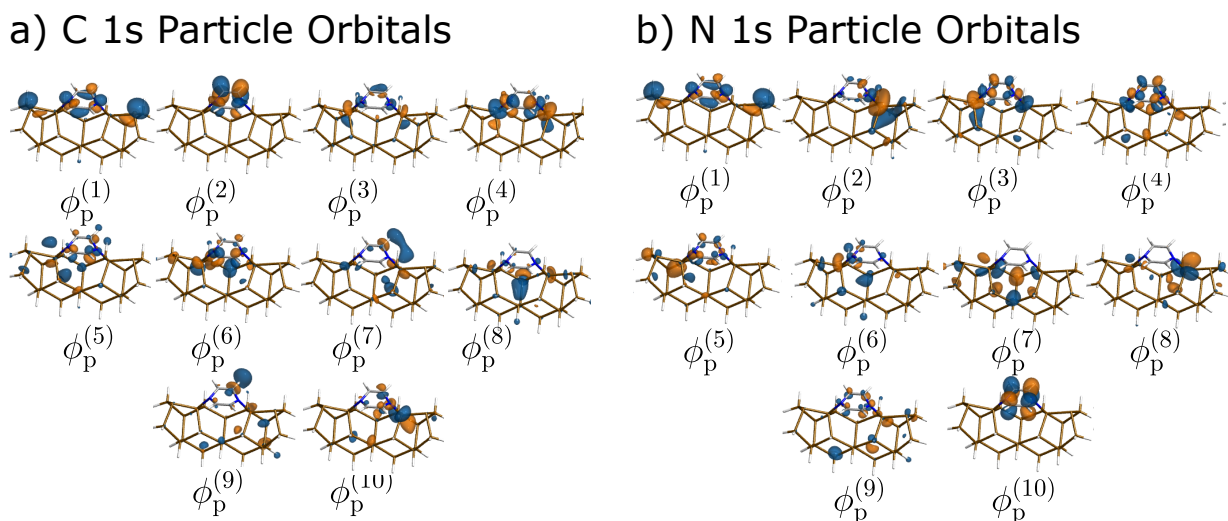


Figure 5.8: Particle orbitals for the core excited states that contribute to the a) carbon and b) nitrogen 1s excitation spectrum of pyrazine ($C_4H_4N_2$) chemisorbed on a $Si_{23}H_{24}$ cluster model of Si(100).

sponds to artifactual excitation to the open Si sites. State 2 is a contribution from the Si-N bond, $\phi_p^{(2)}$ has 39.7% overlap with the $\sigma_{N_4-Si_{12}}^*$ LIVVO. Peak B has contributions from States 3 and 4 both of which are mixed π^*/σ^* states. Similar to the C 1s spectrum, this ratio favors σ^* character due to interaction with the cluster model. The sole contribution to Peak C is a weak transition attributed to the π^* orbitals with 5.2% overlap with $\pi_{C_7-C_8}^*$ and 4.9% overlap with $\pi_{C_1-C_2}^*$.

5.5 Conclusions and Future Work

In this work we have introduced a subspace projection method for orthogonality constrained density functional theory in order to calculate the NEXAFS spectrum of chemisorbed organic molecules. Core excitation energies computed with subspace projected OCDFT are in agreement with those computed using full OCDFT due to the localized nature of these excited states. Throughout the chemisorbed molecules considered here, we notice similar agreement of OCDFT excitation energies to experiment as was observed for gas-phase organic molecules.⁶⁴ We also highlight a powerful tool for analyzing the character of the transitions involved in NEXAFS spectra using

Table 5.3: Calculated (PBE/cc-pVTZ) core excitation energies, oscillator strengths, and LIVVO assignments for the C and N 1s excitation spectrum of pyrazine ($C_4H_4N_2$) chemisorbed on a $Si_{23}H_{24}$ cluster model of Si(100). The LIVVO assignments shown are the two with the highest percentage overlap with the particle orbital. For each calculated transition we report the π^*/σ^* mixing ratio (see text for details) and the spectral feature that this state contributes to (see Figure 5.6).

State	energy (eV)	osc	LIVVO Assignments	π^*/σ^* Mixing	Peak Contribution
C 1s Core Excitations					
1	279.4	0.0002	83.8% $\sigma_{Si_{13}-Si_{14}}^*$	1% $\pi^*/99\%\sigma^*$	
2	283.8	0.0210	73.4% $\pi_{C_7-C_8}^*$ 10.5% $\pi_{C_1-C_2}^*$	86% $\pi^*/24\%\sigma^*$	A
3	284.9	0.0005	11.8% $\sigma_{N_3-Si_{11}}^*$ 10.1% $\sigma_{N_4-Si_{12}}^*$	100% σ^*	B
4	285.3	0.0002	18.5% $\pi_{C_1-C_2}^*$ 13.2% $\sigma_{N_3-Si_{11}}^*$	21% $\pi^*/79\%\sigma^*$	C
5	285.6	0.0006	9.7% $\pi_{C_1-C_2}^*$ 4.0% $\sigma_{C_8-H_{10}}^*$	12% $\pi^*/88\%\sigma^*$	C
6	285.8	0.0005	11.4% $\pi_{C_7-C_8}^*$ 5.7% $\sigma_{N_3-Si_{11}}^*$	14% $\pi^*/86\%\sigma^*$	C
7	286.4	0.0003	10.5% $\sigma_{C_8-H_{10}}^*$ 1.7% $\sigma_{N_3-Si_{11}}^*$	100% σ^*	D
8	286.3	0.0000	2.1% $\pi_{C_7-C_8}^*$ 0.6% $\pi_{C_7-C_8}^*$	3% $\pi^*/97\%\sigma^*+*$	D
9	286.9	0.0005	13.3% $\sigma_{C_8-H_{10}}^*$ 1.6% $\sigma_{C_7-H_9}^*$	2% $\pi^*/98\%\sigma^*$	D
10	287.8	0.0007	12.0% $\sigma_{N_4-Si_{12}}^*$ 3.7% $\sigma_{N_4-C_8}^*$	4% $\pi^*/96\%\sigma^*$	E
N 1s Core Excitations					
1	398.2	0.0026	85.1% $\sigma_{Si_{13}-Si_{14}}^*$ 1.3% $\sigma_{N_4-Si_{12}}^*$	1% $\pi^*/99\%\sigma^*$	A
2	398.4	0.0016	39.7% $\sigma_{N_4-Si_{12}}^*$ 2.1% $\sigma_{N_4-C_8}^*$	2% $\pi^*/98\%\sigma^*$	A
3	399.2	0.0010	12.3% $\sigma_{N_3-Si_{11}}^*$ 7.6% $\pi_{C_7-C_8}^*$	15% $\pi^*/85\%\sigma^*$	B
4	399.5	0.0010	11.0% $\pi_{C_7-C_8}^*$ 6.3% $\sigma_{N_3-Si_{11}}^*$	22% $\pi^*/78\%\sigma^*$	B
5	399.6	0.0005	9.9% $\sigma_{N_3-Si_{11}}^*$ 5.5% $\pi_{C_7-C_8}^*$	12% $\pi^*/88\%\sigma^*$	B
6	399.7	0.0003	5.6% $\sigma_{N_3-Si_{11}}^*$ 2.3% $\pi_{C_7-C_8}^*$	5% $\pi^*/95\%\sigma^*$	B
7	400.2	0.0000	N/A	100% σ^*	
8	400.5	0.0000	N/A	100% σ^*	
9	400.9	0.0003	5.2% $\pi_{C_7-C_8}^*$ 4.9% $\pi_{C_1-C_2}^*$	12% $\pi^*/88\%\sigma^*$	C
10	401.6	0.0000	44.2% $\pi_{C_1-C_2}^*$ 38.9% $\pi_{C_7-C_8}^*$	87% $\pi^*/13\%\sigma^*$	

localized intrinsic valence virtual orbitals. These can assist in more complex cases where simply categorizing an orbital as π or σ is not sufficient, or the molecular orbital character is ambiguous.

We utilize this approach in order to critically analyze the NEXAFS spectrum of pyrazine chemisorbed on the Si(100) surface. Utilizing our subspace projection approach we were able to efficiently target the organic core orbitals of the pyrazine molecule without the need to calculate any lower lying Si core excited states. Applying our LIVVO analysis to assign excited states we were able to track local changes related to the pyrazine molecule upon chemisorption to the surface. Using LIVVOs we were able to quantify the amount of π/σ mixing in each state. Mixing of σ orbitals from the cluster results in fewer states that are purely π^* or purely σ^* with most states showing some degree of mixing. This analysis was most useful in identifying and quantifying differences and changes in intensity. For example, the OCDFT C 1s spectrum for chemisorbed pyrazine showed a large discrepancy between the π and σ manifolds of the spectrum and analysis of the LIVVOs revealed that the π^* state had 73.4% overlap with a single π^* LIVVO while all other states have, 18.5% overlap or less with a single pyrazine LIVVO resulting in significantly weaker transitions.

This subspace projection method in OCDFT lends itself well to future work, the most obvious of which would be the computation of the NEXAFS spectrum of solvated organic molecules. These systems suffer from the same computational hurdle as chemisorbed molecules, where the organic core orbitals of interest are often times higher in energy than those of the solvent molecules. Also for those applications it could be useful to extend this subspace projection formalism to also constrain the choice of the particle orbital to only include unoccupied MOs relevant to the solute. Another potential application is to explore the use of our subspace projection scheme in the context of other excited state methods in order to target specific excitations. Our formulation of this method is general enough to be readily used in this manner.

Appendix A: Implementation of Maximum Subspace Occupation Method

The maximum subspace occupation method presented in this work is a technique that is general enough to be readily applied for other applications. We begin by specifying a subset of atomic orbitals S , a projection onto this subspace can be formally defined as:

$$\hat{\Gamma}_s = \sum_s |\theta_s\rangle\langle\theta_s| \quad (\text{A1})$$

In order to evaluate the occupation of a given molecular orbital within this atomic orbital subspace we must evaluate the expectation value of this operator for the i^{th} MO.

$$\Omega_i = \langle\phi_i|\hat{\Gamma}_s|\phi_i\rangle \quad (\text{A2})$$

$$= \sum_s \langle\phi_i|\theta_s\rangle\langle\theta_s|\phi_i\rangle \quad (\text{A3})$$

We can readily express the MOs of the KS determinant in the AO basis

$$|\phi_i\rangle = \sum_{\mu} C_{i\mu} |\chi_{\mu}\rangle \quad (\text{A4})$$

Using this definition of the MOs

$$\Omega_i = \sum_{\mu\nu} \sum_s \langle\chi_{\mu}|C_{\mu i}^T|\theta_s\rangle\langle\theta_s|C_{i\nu}|\chi_{\nu}\rangle \quad (\text{A5})$$

The subspace basis functions can be defined as an orthonormal transformation of the atomic orbital basis

$$|\theta_s\rangle = \sum_{\mu} X_{\mu s} |\chi_{\mu}\rangle \quad (\text{A6})$$

where X is the orthonormal transformation matrix. we can now plug this definition back into Equation A5 in order to obtain the following expression for the subspace

occupation number.

$$\Omega_i = \sum_{\mu\nu} \sum_s \langle \chi_\mu | C_{\mu i}^T X_{s\nu}^T | \chi_\nu \rangle \langle \chi_\mu | X_{s\mu} C_{i\nu} | \chi_\nu \rangle \quad (\text{A7})$$

$$= \sum_{\mu\nu} \sum_s C_{\mu i}^T \langle \chi_\mu | \chi_\nu \rangle X_{s\nu}^T X_{\mu s} \langle \chi_\mu | \chi_\nu \rangle C_{i\nu} \quad (\text{A8})$$

$$= \mathbf{C}^T \mathbf{s} \mathbf{X} \mathbf{X}^T \mathbf{s} \mathbf{C} \quad (\text{A9})$$

Trivially, if we specify the subspace to include all atomic orbitals, then the subspace occupation number for every MO will be 1.

Bibliography

- [1] Tao, F.; Bernasek, S. L.; Xu, G.-Q. *Chem. Rev.* **2009**, *109*, 3991–4024.
- [2] Bent, S. F. *Surf. Sci.* **2002**, *500*, 879–903.
- [3] Filler, M. A.; Bent, S. F. *Prog. Surf. Sci.* **2003**, *73*, 1–56.
- [4] Wolkow, R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 413–441.
- [5] Karthäuser, S. *J. Phys.: Condens. Matter* **2011**, *23*, 013001.
- [6] Mentovich, E.; Kalifa, I.; Tsukernik, A.; Caster, A.; Rosenberg-Shraga, N.; Marom, H.; Gozin, M.; Richter, S. *Small* **2008**, *4*, 55–58.
- [7] Joachim, C.; Gimzewski, J. K.; Aviram, A. *Nature* **2000**, *408*, 541–548.
- [8] Mayne, A. J.; Lastapis, M.; Baffou, G.; Soukiassian, L.; Comtet, G.; Hellner, L.; Dujardin, G. *Phys. Rev. B* **2004**, *69*, 045409.
- [9] Lu, X. *J. Am. Chem. Soc.* **2003**, *125*, 6384–6385.
- [10] Taguchi, Y.; Fujisawa, M.; Takaoka, T.; Okada, T.; Nishijima, M. *J. Chem. Phys.* **1991**, *95*, 6870–6876.
- [11] Sinniah, K.; Sherman, M. G.; Lewis, L. B.; Weinberg, W. H.; Yates, J. T.; Janda, K. C. *Phys. Rev. Lett.* **1989**, *62*, 567–570.
- [12] Zhao, J.; Noffke, B. W.; Raghavachari, K.; Tepyakov, A. V. *J. Phys. Chem. C* **2017**, *121*, 7208–7213.
- [13] Akagi, K.; Yoshinobu, J. *Surf. Sci.* **2016**, *652*, 304–311.
- [14] Koeny, R.; Doren, D. J. *Surf. Sci.* **1998**, *417*, 169–188.
- [15] Lu, X.; Xu, X.; Wu, J.; Wang, N.; Zhang, Q. *New J. Chem.* **2002**, *26*, 160–164.
- [16] Qu, Y.-Q.; Han, K.-L. *J. Phys. Chem. B* **2004**, *108*, 8305–8310.
- [17] Tao, F.; Wang, Z. H.; Xu, G. Q. *J. Phys. Chem. B* **2002**, *106*, 3557–3563.

- [18] Kong, M. J.; Teplyakov, A. V.; Lyubovitsky, J. G.; Bent, S. F. *Surf. Sci.* **1998**, *411*, 286–293.
- [19] Wang, G. T.; Mui, C.; Tannaci, J. F.; Filler, M. A.; Musgrave, C. B.; Bent, S. F. *J. Phys. Chem. B* **2003**, *107*, 4982–4996.
- [20] Romeo, M.; Balducci, G.; Stener, M.; Fronzoni, G. *J. Phys. Chem. C* **2014**, *118*, 1049–1061.
- [21] Tao, F.; Qiao, M. H.; Wang, Z. H.; Xu, G. Q. *J. Phys. Chem. B* **2003**, *107*, 6384–6390.
- [22] Miwa, J. A.; Eves, B. J.; Rosei, F.; Lopinski, G. P. *J. Phys. Chem. B* **2005**, *109*, 20055–20059.
- [23] Weier, D.; Lühr, T.; Beimborn, A.; Schönbohm, F.; Döring, S.; Berges, U.; Westphal, C. *Surf. Sci.* **2011**, *605*, 1784–1790.
- [24] Ardalan, P.; Dupont, G.; Musgrave, C. B. *J. Phys. Chem. C* **2011**, *115*, 7477–7486.
- [25] Chatterjee, A.; Zhang, L.; Leung, K. T. *Langmuir* **2013**, *29*, 9369–9377.
- [26] Jung, S. C.; Kang, M. H. *Phys. Rev. B* **2009**, *80*, 235312.
- [27] Huang, H. G.; Huang, J. Y.; Ning, Y. S.; Xu, G. Q. *J. Chem. Phys.* **2004**, *121*, 4820–4825.
- [28] Lee, H.-K.; Park, J.; Kim, I.; Kim, H.-D.; Park, B.-G.; Shin, H.-J.; Lee, I.-J.; Singh, A. P.; Thakur, A.; Kim, J.-Y. *J. Phys. Chem. C* **2012**, *116*, 722–725.
- [29] Lu, X.; Lin, M. *Int. Rev. Phys. Chem.* **2002**, *21*, 137–184.
- [30] Ng, W. K. H.; Sun, S. T.; Liu, J. W.; Liu, Z. F. *J. Phys. Chem. C* **2013**, *117*, 15749–15753.
- [31] Omiya, T.; Yokohara, H.; Shimomura, M. *J. Phys. Chem. C* **2012**, *116*, 9980–9984.
- [32] Shimomura, M. *Asian J. Chem.* **2013**, *25*, S35.
- [33] Stöhr, J. *NEXAFS Spectroscopy*; Springer: Berlin, 1992; Vol. 25.
- [34] Penner-Hahn, J. E. *Coord. Chem. Rev.* **1999**, *190192*, 1101–1123.
- [35] Garino, C.; Borfecchia, E.; Gobetto, R.; van Bokhoven, J. A.; Lamberti, C. *Coord. Chem. Rev.* **2014**, *277278*, 130–186.
- [36] Tourillon, G.; Mahatsekake, C.; Andrieu, C.; Williams, G. P.; Garrett, R. F.; Braun, W. *Surf. Sci.* **1988**, *201*, 171–184.
- [37] Rosenberg, R. A.; Love, P. J.; Rehn, V. *Phys. Rev. B* **1986**, *33*, 4034–4037.
- [38] Shimoyama, I.; Wu, G.; Sekiguchi, T.; Baba, Y. *Phys. Rev. B* **2000**, *62*, R6053–R6056.
- [39] Stöhr, J.; Outka, D. A. *Phys. Rev. B* **1987**, *36*, 7891–7905.
- [40] This L stands for Langmuir, it is a standard unit to quantify exposure of a surface to a gaseous sample. The quantity is obtained by multiplying the pressure of the gas by the time of exposure.
- [41] Triguero, L.; Pettersson, L. G. M.; Ågren, H. *Phys. Rev. B* **1998**, *58*, 8097–8110.
- [42] Besley, N. A.; Noble, A. *J. Phys. Chem. C* **2007**, *111*, 3333–3340.
- [43] Fronzoni, G.; Balducci, G.; De Francesco, R.; Romeo, M.; Stener, M. *J. Phys. Chem. C* **2012**, *116*, 18910–18919.

- [44] De Francesco, R.; Stener, M.; Fronzoni, G. *J. Phys. Chem. C* **2007**, *111*, 13554–13563.
- [45] Coustel, R.; Carniato, S.; Rochet, F.; Witkowski, N. *Phys. Rev. B* **2012**, *85*, 035323.
- [46] Rangan, S.; Gallet, J.-J.; Bournel, F.; Kubsy, S.; Le Guen, K.; Dufour, G.; Rochet, F.; Sirotti, F.; Carniato, S.; Ilakovac, V. *Phys. Rev. B* **2005**, *71*, 165318.
- [47] Rangan, S.; Bournel, F.; Gallet, J.-J.; Kubsy, S.; Le Guen, K.; Dufour, G.; Rochet, F.; Sirotti, F.; Carniato, S.; Ilakovac, V. *Phys. Rev. B* **2005**, *71*, 165319.
- [48] Fransson, T.; Coriani, S.; Christiansen, O.; Norman, P. *J. Chem. Phys.* **2013**, *138*, 124311.
- [49] Besley, N. A. *Chem. Phys. Lett.* **2012**, *542*, 42–46.
- [50] Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. *Phys. Rev. A* **2012**, *85*, 022507.
- [51] Coriani, S.; Fransson, T.; Christiansen, O.; Norman, P. *J. Chem. Theory Comput.* **2012**, *8*, 1616–1628.
- [52] Myhre, R. H.; Coriani, S.; Koch, H. *J. Chem. Theory Comput.* **2016**, *12*, 2633–2643.
- [53] Maganas, D.; Kristiansen, P.; Duda, L.-C.; Knop-Gericke, A.; DeBeer, S.; Schlögl, R.; Neese, F. *J. Phys. Chem. C* **2014**, *118*, 20163–20175.
- [54] Maganas, D.; Roemelt, M.; Weyhermuller, T.; Blume, R.; Havecker, M.; Knop-Gericke, A.; DeBeer, S.; Schlögl, R.; Neese, F. *Phys. Chem. Chem. Phys.* **2014**, *16*, 264–276.
- [55] Grimme, S. *Chem. Phys. Lett.* **1996**, *259*, 128–137.
- [56] Wenzel, J.; Wormit, M.; Dreuw, A. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.
- [57] Wenzel, J.; Holzer, A.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2015**, *142*, 214104.
- [58] Wenzel, J.; Dreuw, A. *J. Chem. Theory Comput.* **2016**, *12*, 1314–1330.
- [59] Imamura, Y.; Nakai, H. *Chem. Phys. Lett.* **2006**, *419*, 297–303.
- [60] Lestrangle, P. J.; Nguyen, P. D.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.
- [61] Asmuruf, F.; Besley, N. *Chem. Phys. Lett.* **2008**, *463*, 267–271.
- [62] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012**, *8*, 3284–3292.
- [63] Evangelista, F. A.; Shushkov, P.; Tully, J. C. *J. Phys. Chem. A* **2013**, *117*, 7378–7392.
- [64] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, 14360–14374.
- [65] Verma, P.; Derricotte, W. D.; Evangelista, F. A. *J. Chem. Theory Comput.* **2016**, *12*, 144–156.
- [66] Stener, M.; Fronzoni, G.; de Simone, M. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- [67] Besley, N. A.; Asmuruf, F. A. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024.
- [68] Peng, B.; Lestrangle, P. J.; Goings, J. J.; Caricato, M.; Li, X. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.
- [69] Derricotte, W. D.; Evangelista, F. A. *Submitted to J. Chem. Theory Comput.* **2017**,

- [70] Ayers, P. W.; Levy, M.; Nagy, A. *Phys. Rev. A* **2012**, *85*, 042518.
- [71] Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.
- [72] Knizia, G. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- [73] Schmidt, M. W.; Hull, E. A.; Windus, T. L. *J. Phys. Chem. A* **2015**, *119*, 10408–10427.
- [74] Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- [75] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- [76] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [77] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [78] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [79] Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223–1229.
- [80] Hariharan, P. C.; Pople, J. A. *Theoret. Chim. Acta* **1973**, *28*, 213–222.
- [81] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- [82] Turney, J. M. et al. *WIREs Comput Mol Sci* **2012**, *2*, 556–565.
- [83] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [84] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- [85] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- [86] Ma, Y.; Sette, F.; Meigs, G.; Modesti, S.; Chen, C. T. *Phys. Rev. Lett.* **1989**, *63*, 2044–2047.
- [87] Carravetta, V.; Ågren, H. *J. Phys. Chem. A* **2013**, *117*, 6798–6802.
- [88] Vall-llosera, G.; Gao, B.; Kivimäki, A.; Coreno, M.; Álvarez Ruiz, J.; de Simone, M.; Ågren, H.; Rachlew, E. *J. Chem. Phys.* **2008**, *128*, 044316.

Chapter 6

Concluding Remarks and Outlook

“Mathematical analysis is as extensive as nature itself; it defines all perceptible relations, measures times, spaces, forces, temperatures; this difficult science is formed slowly, but it preserves every principle which it has once acquired; it grows and strengthens itself incessantly in the midst of the many variations and errors of the human mind.”

Joseph Fourier

In this dissertation I have presented novel developments of orthogonality constrained density functional theory (OCDFT) for the treatment of electronic core-excited states. These new features aim to craft a robust computational approach for core-excited states that can provide accurate energies, properties, detailed orbital assignments, and versatility to handle unique chemical environments.

The initial study focuses on introducing two unique features to OCDFT. First the OCDFT eigenvalue equation is augmented such that the targeted electronic transitions are those involving hole orbitals with the *lowest* eigenvalues, thus targeting the core region of the excitation spectrum. This extension, however, only allows for the computation of a single core-excited state of a given symmetry. In order to allow for full spectral simulations, the constrained multiple hole/particle (CMHP) algorithm is introduced for calculating multiple excited state solutions. Benchmark calculations were performed on a series of 23 core excitations involving core first-row elements and 17 from second-row elements. At the B3LYP/def2-QZVP level of

theory, OCDFT consistently outperformed time-dependent density functional theory (TDDFT). [W.D. Derricotte and F. A. Evangelista, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14360] Comparing the first-row excitations to gas-phase experimental values, OCDFT produces a mean absolute error (MAE) of 0.4 eV, while TDDFT performs significantly worse with a MAE of 11.6 eV. For the second-row excitations, OCDFT and TDDFT produce MAEs of 1.6 and 31.6 eV respectively. The two theories were formally compared using a model system of two electrons in two orbitals, in order to investigate the source of error. Using this analysis, the error can be related to nonlocal Coulomb repulsion integrals that arise in the expressions of TDDFT but not in OCDFT, providing an incorrect description of the hole/particle physics of core excitations. In the case of nucleobase molecules thymine and adenine, OCDFT was also shown to produce full NEXAFS spectral simulations that provide a good comparison to experimental results. Experimental peak positions calculated with OCDFT have an average error of 0.3 eV for thymine and 0.1 eV for adenine, with the calculated oscillator strengths providing an accurate representation of experimental peak intensities.

The explicit treatment of relativistic effects for core-excited states was explored next through the implementation of the spin-free exact-two-component (X2C) relativistic Hamiltonian. This scalar relativistic treatment was coupled with OCDFT in order to investigate the impact of relativistic effects on core excitation energies. Our results show that for first-row core excitations X2C-OCDFT exhibits a slight improvement of 0.2 eV in the MAE compared to a non-relativistic treatment, confirming that scalar relativistic effects are not a large source of error for first-row core excitations. However, for second-row core excitations X2C-OCDFT shows a 10 eV improvement in the MAE compared to a non-relativistic treatment. Full treatment of relativistic effects via the X2C Hamiltonian made it possible to properly treat core excitations of transition metal complexes. This was shown by simulating the

Ti K-edge of three Ti containing complexes TiCl_4 , TiCpCl_3 , and TiCp_2Cl_2 (where “Cp” represents cyclopentadiene). The Ti 1s core excitation energies were calculated to within approximately 5.0 eV of the experimental spectrum. This represents an improvement of 30 eV over a nonrelativistic OCDFT treatment and a 70 eV improvement over a relativistic TDDFT treatment. The atomic contributions to the pre-edge features of the Ti K-edge were critically analyzed using an atomic decomposition of the OCDFT transition dipole moment (TDM) along with a population analysis of the natural atomic orbitals (NAO). The largest atomic contributions to the total dipole are shown to be Ti s \rightarrow Ti p transitions and Ti s \rightarrow Cl p transitions with these two contributions accounting for roughly 80% of the total dipole in all pre-edge transitions. Meanwhile, the largest change in NAO population occurs in the Ti d orbitals, with a change in Ti d population of greater than 1.0 for all pre-edge transitions. This is consistent with an interpretation that the pre-edge features of these transition metal complexes correspond to 1s \rightarrow 3d transitions that are forbidden by selection rules, yet gain intensity through mixing with Ti and ligand p orbitals.

In an effort to provide more robust assignments of core-excited states in OCDFT, a new orbital representation is presented, known as the localized intrinsic valence virtual orbitals (LIVVOs). In this approach, the LIVVOs are derived and classified based on the intrinsic atomic orbitals (IAOs). The LIVVOs are then used to give a detailed classification of the OCDFT particle orbitals by quantifying the localized contributions. This method of classification was tested on ethanethiol and benzenethiol molecules. In both molecules, the first two excited states correspond to particle orbitals that span both thiol bond (S-H) and the adjacent S-C bond. It is difficult from visual inspection of the particle orbitals to discern the contribution of the thiol bond. Using our analysis we were able to show that the first state is dominated by the σ_{S-H}^* LIVVO with a 58% contribution in ethanethiol and a 61% contribution in benzenethiol. Meanwhile the second state was more localized along the σ_{S-C}^* LIVVO

with a 57% and 53% contribution from ethanethiol and benzenethiol, respectively. Further tests on the water monomer and dimer showed the LIVVO analysis useful in quantifying differences in the excited states due to changes in the molecular environment. The particle orbitals for the water monomer showed equal contributions from both σ_{O-H}^* LIVVOs. When simulating the 1s core excitation spectrum of the accepting O atom in the water dimer, both σ_{O-H}^* LIVVOs still show equal contribution, suggesting that the spectrum is unaffected by accepting a hydrogen bond. However, in the spectrum of the donating O atom, 52% of the electron density in the particle orbital of the first core-excited is localized on the LIVVO with only 16% of the electron density lying along the OH bond participating in the hydrogen bond. Our LIVVO analysis was able to highlight and quantify the significant effect that donating a hydrogen bond has on the local O core excitation spectrum.

Lastly I presented an extension to the CMHP algorithm in OCDFT that allows for the targeting of specific hole orbitals based on their occupation in an atomic orbital subspace. This maximum subspace occupation (MSO) method allows for targeting of core orbitals in situations where they are not the lowest energy eigenvalues in the spectrum. As a case study, the C and N K-edges of a pyrazine ($C_4H_4N_2$) molecule chemisorbed to a Si(100) surface was studied. MSO-OCDFT was able to target the organic core orbitals without first calculating core excitations from the lower energy Si 1s orbitals. This alleviates a computational difficulty that would have made this calculation infeasible in the original CMHP algorithm. Subsequent LIVVO analysis of the pyrazine spectrum reveals interesting changes upon chemisorption to the surface. For example the first state in the C 1s spectrum maintains similar localization about the π^* LIVVO of the C=C double bond to the gas phase spectrum. In the gas phase the particle orbital has 78.2% π_{C-C}^* LIVVO, when chemisorbed, it has 73.4% overlap with this LIVVO. As a result the intensity is identical in both cases ($f_{abs} = 0.02$). Meanwhile the localization of all states in the σ^* manifold drop by 35% in all cases.

Leading to a dramatic difference in intensity between the π^* and σ^* manifolds in the chemisorbed spectrum.

This dissertation establishes OCDFT as an accurate and cost effective method for the computation of core-excited states. The theory has been significantly enhanced with new features that allow for the treatment of a wide range of different chemical systems including chemisorbed organic molecules and transition metal complexes. Further work on the theory is required in order to encompass a more complete treatment of the entire field of X-ray spectroscopy. While the work presented here was limited to K-edge applications, extensions to treat L-edge spectroscopy are possible. This would require pairing OCDFT with a higher-order relativistic treatment to account for the spin-orbit splitting that becomes a factor in 2p core-excitations. Another desirable extension would be a method to treat vibrational effects within the core spectrum, resulting in a more realistic representation of the spectral resonances. One can imagine coupling OCDFT with a discrete variable representation (DVR) approach for the treatment of vibrations. Lastly, OCDFT could be extended for the treatment of X-ray emission spectroscopy. This problem would require the calculation of an approximate intermediate state in which a core hole is introduced. This core hole would then be the LUMO for all subsequent excited state calculations and thus the lowest energy solutions to the OCDFT eigenvalue equation would be the x-ray emission energies.