THE DEVELOPMENT AND APPLICATIONS OF A NOVEL SOFTWARE PACKAGE FOR THE ANALYSIS OF ELECTRON PAIR BEHAVIOUR

A Thesis

Submitted to the Graduate Faculty

in Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy

Molecular and Macromolecular Sciences

Department of Chemistry

Faculty of Science

University of Prince Edward Island

Adam J. Proud

Charlottetown, Prince Edward Island

August 2018

University of Prince Edward Island
Faculty of Science
Charlottetown

CERTIFICATION OF THESIS WORK

We, the undersigned, certify that Adam Jonathan Proud, candidate for the degree of Doctor of Philosophy has presented a thesis with the following title: The development and applications of a novel software package for the analysis of electron pair behaviour, that the thesis is acceptable in form and content, and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate through an oral examination held on August 28 2018.

Examiners: Dr. Jason Pearson [redacted]

Dr. Brian Wagner [redacted]

Dr. Nasser Saad [redacted]

Dr. Joshua Hollett [redacted]

Dr. Russell Kerr - Chair [redacted]

Date AUG 28, 2018

THESIS/DISSERTATION NON-EXCLUSIVE LICENSE

Family Name: Proud	Given Name, Middle Name (if applicable): Adam Jonathan
Full Name of University: University of Prince Edward Island	
Faculty, Department, School: Faculty of Science	
Degree for which thesis/dissertation was presented: Ph.D. in Molecular and Macromolecular Sciences	Date Degree Awarded: August 28, 2018
Thesis/dissertation Title: The Development and Applications of a Novel Software Package for the Analysis of Electron Pair Behaviour	
Date of Birth. It is **optional** to supply your date of birth. If you choose to do so please note that the information will be included in the bibliographic record for your thesis/dissertation.	

In consideration of my University making my thesis/dissertation available to interested persons, I,

Adam Jonathan Proud

hereby grant a non-exclusive, for the full term of copyright protection, license to my University, The university of Prince Edward Island:

(a) to archive, preserve, produce, reproduce, publish, communicate, convert into any format, and to make available in print or online by telecommunication to the public for non-commercial purposes;
(b) to sub-license to Library and Archives Canada any of the acts mentioned in paragraph (a).

I undertake to submit my thesis/dissertation, through my University, to Library and Archives Canada. Any abstract submitted with the thesis/dissertation will be considered to form part of the thesis/dissertation.

I represent that my thesis/dissertation is my original work, does not infringe any rights of others, including privacy rights, and that I have the right to make the grant conferred by this non-exclusive license.

If third party copyrighted material was included in my thesis/dissertation for which, under the terms of the *Copyright Act*, written permission from the copyright owners is required I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection

I retain copyright ownership and moral rights in my thesis/dissertation, and may deal with the copyright in my thesis/dissertation, in any way consistent with rights granted by me to my University in this non-exclusive license.

I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis/dissertation of the rights granted by me to my University in this non-exclusive license.

Signature	Date
[redacted]	September 24, 2018

Abstract

The “Lewis Pair” is a ubiquitous phenomenon in chemistry and is often used as an intuitive construct to predict and rationalize chemical structure and behaviour. Concepts from the very general Valence Shell Electron Pair Repulsion (VSEPR) model to the most esoteric reaction mechanism routinely rely on the notion that electrons tend to exist in pairs and that these pairs can be thought of as being localized to a particular region of space. It is this localization that allows one to intuit how these pairs might behave, generally speaking, so that reasonable predictions may be made regarding molecular structure, intermolecular interactions, property trends, and reaction mechanisms, etc. Of course, it is rather unfortunate that the Lewis model is entirely qualitative and yields no information regarding how any specific electron pair is distributed.

Herein, we demonstrate a novel electronic structure analysis technique that predicts and analyzes precise quantitative details about the relative and absolute distribution of individual electron pairs. This Single Electron Pair Distribution Analysis (SEPDA) reveals important quantitative details about the distribution of the well-known Lewis pairs, such as how they are distributed in space and how their relative velocities change in various chemical contexts. We show that such an analysis may be used to quantify and classify a wide range of interactions including chemical bonding and non-covalent interactions. The nature of non-covalent interactions (as well as indications of their strength) may also be gleaned from such distributions and SEPDA can be used as an important tool to differentiate between interaction

types. Specifically, we use the SEPDA package to analyze covalent compounds in terms of their electron behaviour and electronegativity properties, as well as non-covalent interactions such as hydrogen bonding.

We anticipate that SEPDA will be of broad utility in a wide variety of chemical contexts because it affords a very detailed, visual and intuitive analysis technique that is generally applicable. The development of a user-friendly, publicly available software package should only further prove the wide applicability and significance of SEPDA.

Acknowledgements

It has been a long road for me at the University of Prince Edward Island having completed an Honour's thesis, a Master's thesis and now a Ph.D. thesis. I definitely could not have accomplished this without the help of countless people.

Dr. Jason Pearson has been my mentor and supervisor since for the past 8 years. Under his tutelage, I have grown exceptionally as a researcher and I cannot thank him enough for everything he has done for me over the course of my two graduate degrees. For this, Jason, I sincerely thank you.

I would also like to thank my graduate supervisory committee: Dr. Barry Linkletter, Dr. Sheldon Opps, and Dr. Brian Wagner. You all provided me with significant feedback and helped guide me through my doctoral studies. I greatly appreciate all of the help that each of you has provided me over the years.

Throughout my time in the Pearson group, many students have come and gone. I would like to thank you all (Brendan Sheppard, Zosia Zielinski, Dalton Mackenzie, Dylan Hennessey, Meagan Oakley, Ellen O'Connor, Simon Sirois-Lecain, Andrew Cameron, Qammar Almas, Ben Keefe, Trevor Profitt, Mat Larade, and Trevor Profitt) for being friends, for being supportive, and always being there if I ever needed anything. In terms of my actual research, I would specifically like to thank Brendan, Zosia, and Dalton for contributions to this research.

The Chemistry Department at UPEI is small, but very strong. Faculty, staff, and students make up the best department on campus. I cannot begin to express my thanks to everyone in this department for all that they have done for me throughout

my tenure at UPEI. Specifically, I would like to thank Dr. Rabin Bissessur for being a strong mentor to me over the years from a teacher, to a supervisor, and finally as a strong role model and colleague.

Much of this research was made possible my existing recurrence relations developed by Dr. Ajit Thakkar and by Dr. Peter Gill and Dr. Joshua Hollett. I would like to thank each of these scientists for insightful discussions regarding their research, and especially, Dr. Hollett, for being a part of my thesis defence.

None of this research would have been possible without funding from UPEI, the Natural Sciences and Engineering Research Council of Canada, and ACENet. Further, the millions of hours of computational time required for this project were provided by Compute Canada (specifically ACENet and WestGrid).

Finally, and possibly most importantly, I would like to thank my friends both at UPEI and outside, my family, specifically, Ryan and Adrienne, my parents, Mike and Anne, and my girlfriend Courtney for all of their love and support. These past 4 years have been some of the most difficult of my life. I don't know where I would be without you all in my life. Thank you for always being there for me, especially when I needed you the most.

To Nana Lawlor

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

AO	Atomic orbital
CMO	Canonical molecular orbital
DFT	Density functional theory
ER	Edmiston-Ruedenberg
HF	Hartree-Fock
LCAO	Linear combination of atomic orbitals
LMO	Localized molecular orbital
MO	Molecular orbital
MPPT	Møller-Plesset perturbation theory
NBO	Natural bond orbital
RHF	Restricted Hartree-Fock
SCF	Self consistent field
UHF	Unrestricted Hartree-Fock

Symbols

Ψ	Molecular wavefunction
ψ	Spatial molecular orbital
χ	Spin molecular orbital
ϕ	Basis function/atomic orbital
$\hat{H}$	Hamiltonian operator
$\hat{T}$	Kinetic energy operator
$\hat{V}$	Potential energy operator
$\hat{H}^{elec}$	Electronic Hamiltonian operator
$\hat{H}^{core}$	Core Hamiltonian operator
$\hat{H}_2$	Two-electron Hamiltonian operator
$\hat{P}_{ij}$	Permutation operator

$\hat{J}$	Coulomb operator
$\hat{K}$	Exchange operator
∇	Gradient operator ($\nabla \equiv \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$)
∇^2	Laplacian operator ($\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$)
$\boldsymbol{r_i}$	Position vector for electron i
$\boldsymbol{p_i}$	Momentum vector for electron i
$\boldsymbol{x}_i$	Position-spin vector for electron i
r_{ij}	Distance between electrons i and j
R_{AB}	Distance between nuclei A and B
Z_A	Atomic number of nucleus A
N	Number of electrons
M	Number of nuclei
K	Number of basis functions
$S_{\mu\nu}$	Element μ, ν of the overlap matrix
$F_{\mu\nu}$	Element μ, ν of the Fock matrix
$P_{\mu\nu}$	Element μ, ν of the charge density matrix
$\boldsymbol{F}$	Fock matrix
$\boldsymbol{C}$	MO coefficient matrix
$\boldsymbol{S}$	Overlap matrix
$\boldsymbol{X}$	Overlap transformation matrix
$\boldsymbol{E}$	MO energy matrix
E	Energy
$c_{\mu i}$	AO coefficient μ for MO i
d_i	Basis function contraction coefficient
$\alpha(\omega)$	Spin-up function
$\beta(\omega)$	Spin-down function
$\delta(x)$	Dirac delta function
$\rho(\boldsymbol{r})$	Electron density
$\rho(\boldsymbol{r_1}, \boldsymbol{r_2})$	Two-electron density or pair density
$P(u)$	Position intracule density
$M(v)$	Momentum intracule density
$X(x)$	Posmom intracule density
$E(R)$	Position extracule density
$E(\boldsymbol{R})$	3-D Position extracule density
$E(P)$	Momentum extracule density
u	$P(u)$ coordinate, equal to $\lvert\boldsymbol{r_1} - \boldsymbol{r_2}\rvert$
v	$M(v)$ coordinate, equal to $\lvert\boldsymbol{p_1} - \boldsymbol{p_2}\rvert$
x	$X(x)$ coordinate, equal to $\lvert\boldsymbol{u} \cdot \boldsymbol{v}\rvert$
R	$E(R)$ coordinate, equal to $\lvert\boldsymbol{r_1} + \boldsymbol{r_2}\rvert/2$
$\boldsymbol{R}$	$E(\boldsymbol{R})$ coordinate vector
P	$E(P)$ coordinate, equal to $\lvert\boldsymbol{p_1} + \boldsymbol{p_2}\rvert/2$
$\mathrm{d}\Omega_i$	Integration over all angular components of vector i

Chapter 1

Introduction

It was long assumed that Newton's laws of motion[5] governed the motion of everything. By knowing the initial position, $\boldsymbol{r}_0$, and velocity, $\boldsymbol{v}_0$, of an object along with its mass, m, and the net force acting on it, $\boldsymbol{F}$, one could determine the future position, $\boldsymbol{r}(t)$ of said object at any time, t, through the equation

$$\boldsymbol{r}(t) = \frac{\boldsymbol{F}}{2m}t^2 + \boldsymbol{v}_0 t + \boldsymbol{r_0} \tag{1.1}$$

This relationship is easily obtained through integration from the more common form of Newton's second law of motion: $\boldsymbol{F} = m\boldsymbol{a}$. Obviously not all situations are as simple as those that are adequately described by this equation since most objects do not have constant forces acting on them. Complicating the issue even further is the many-body problem. Consider the problem of planetary motion. While the planets in a given solar system typically revolve around the sun based on the gravitational force, one also has to consider the interaction between the planets themselves. It

is this many-body interaction that greatly complicates the situation, and it is one that we are faced with in all chemical species with the exception of one-electron, one-nucleus systems (e.g. the hydrogen atom).[6–8]

The other major issue with applying Newton's laws of motion to chemical systems can be explained by the Heisenberg uncertainty principle.[9] It states that "The more precisely the position is determined, the less precisely the momentum is known in this instant, and vice versa" (translation by the American Institute of Physics). More simply, one cannot know both the momentum and position of a particle, simultaneously. There is an uncertainty principle associated with these terms. Mathematically, Heisenberg's uncertainty principle is given by

$$\Delta x \Delta p \geq \hbar/2 \tag{1.2}$$

where Δx and Δp denote the uncertainty in the position and momentum, respectively. This uncertainty applies to all objects, even those that behave by the laws of classical mechanics. However, when one considers the size of the uncertainty (on the order of $10^{-35}\ kg \cdot m^2 \cdot s^{-1}$), the uncertainty in either the position or momentum of an object even on the nanogram scale would be negligible. When, however, we consider the size of sub atomic particles with masses on the order of 10^{-31} kg, the uncertainty in momentum is rather substantial. Instead, these microscopic systems must be treated with quantum mechanics which is probabilistic in nature.

1.1 Quantum Mechanics

Much like Newton's laws of classical mechanics are instrumental in understanding the motion of objects we see in our everyday lives, the Schrödinger equation[10] is at the heart of quantum mechanics and understanding systems at the electronic level. The time-dependent Schrödinger equation is given by

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle \tag{1.3}$$

where $|\Psi\rangle$ is the molecular wavefunction and $\hat{H}$ is known as the Hamiltonian operator. The wavefunction contains all of the information needed to completely describe a chemical system, while the Hamiltonian is the operator, which when applied to the wavefunction, yields the total energy of the system. Fortunately, for nearly all applications in chemistry, the time-dependent Schrödinger equation is unnecessary and we can simplify to the time-independent form, which from this point forth, we will simply refer to as the Schrödinger equation.[6] This form is written as

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{1.4}$$

The Hamiltonian operator consists of two types of operators, those pertaining to potential energy and those pertaining to kinetic energy. The Hamiltonian is given by[6]

$$\hat{H} \equiv -\frac{\hbar^2}{2m_e}\sum_{i=1}^{N}\nabla_i^2 - \sum_{A=1}^{M}\frac{\hbar^2}{2M_A}\nabla_A^2 + V(\boldsymbol{r}) \tag{1.5}$$

where the first two terms correspond to the kinetic energy of the N electrons which have mass m and the M nuclei with mass M_A, respectively. The third term is the potential energy operator; however, before proceeding with its definition, it is much simpler to introduce a new set of units, known as atomic units. Atomic units are frequently used in computational chemistry to express everything from lengths to masses and energies. By using atomic units[11] one can greatly simplify the mathematical form of the Hamiltonian operator, especially the terms pertaining to potential energy. The list of atomic units relevant in this thesis are given below in Table 1.1.

Using this new set of units, we can define the potential energy operator as

$$V(r) \equiv \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{1.6}$$

with the first and third terms defining the electron-electron and nuclear-nuclear repulsions, respectively, which are dependent on the interelectron, r_{ij} and internuclear, R_{AB}, distances. The nuclear-nuclear repulsion is also dependent on the nuclear charges of atoms A and B (Z_A and Z_B). The second term describes the

Table 1.1: Definition of atomic units

Measure	Unit	Value in atomic units	Value in SI Units
Length	a_0	1 bohr	5.2918×10^{-11} m
Mass	m_e	1	9.1095×10^{-31} kg
Charge	e	1	1.6022×10^{-19} C
Energy	E	1 hartee (E_h)	4.3498×10^{-18} J
	E	1 hartee (E_h)	27.211 eV
	E	1 hartee (E_h)	627.51 kcal mol^{-1}
Angular momentum	$\hbar$	1	1.0546×10^{-34} J s
Vacuum permittivity	$4\pi\varepsilon_0$	1	1.113×10^{-10} C^2 J^{-1} m^{-1}

electron-nuclear attraction which is dependent on charge of nucleus A as well as the distance between the nucleus and electron i, r_{iA}. Combining this equation with (1.5) yields the expression for the total Hamiltonian

$$\hat{H} \equiv -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{1.7}$$

Thus, in its full form, the Schrödinger equation is given by

$$\left[-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 + \left(\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}\right)\right]\Psi(\mathbf{r},\mathbf{R}) = E\Psi(\mathbf{r},\mathbf{R}) \tag{1.8}$$

For the sake of clarity, the terms in the Hamiltonian are colour coordinated with the visual representation of the terms in Figure 1.1. In this form, the Schrödinger equation looks quite daunting. However, the equation can be simplified further by invoking the Born-Oppenheimer approximation.

1.2 Born-Oppenheimer Approximation

Consider that the mass of a proton is approximately 1840 times that of an electron. Even for this smallest of atomic systems, the electrons would be moving far faster than the nuclei. This only amplifies as we study systems containing nuclei larger than that of hydrogen. Thus, to a very good approximation, it can be assumed that the electrons in a molecule move in a field of fixed nuclei. This is known as the Born-Oppenheimer approximation.[12] Physicists tend to use a slightly more complicated

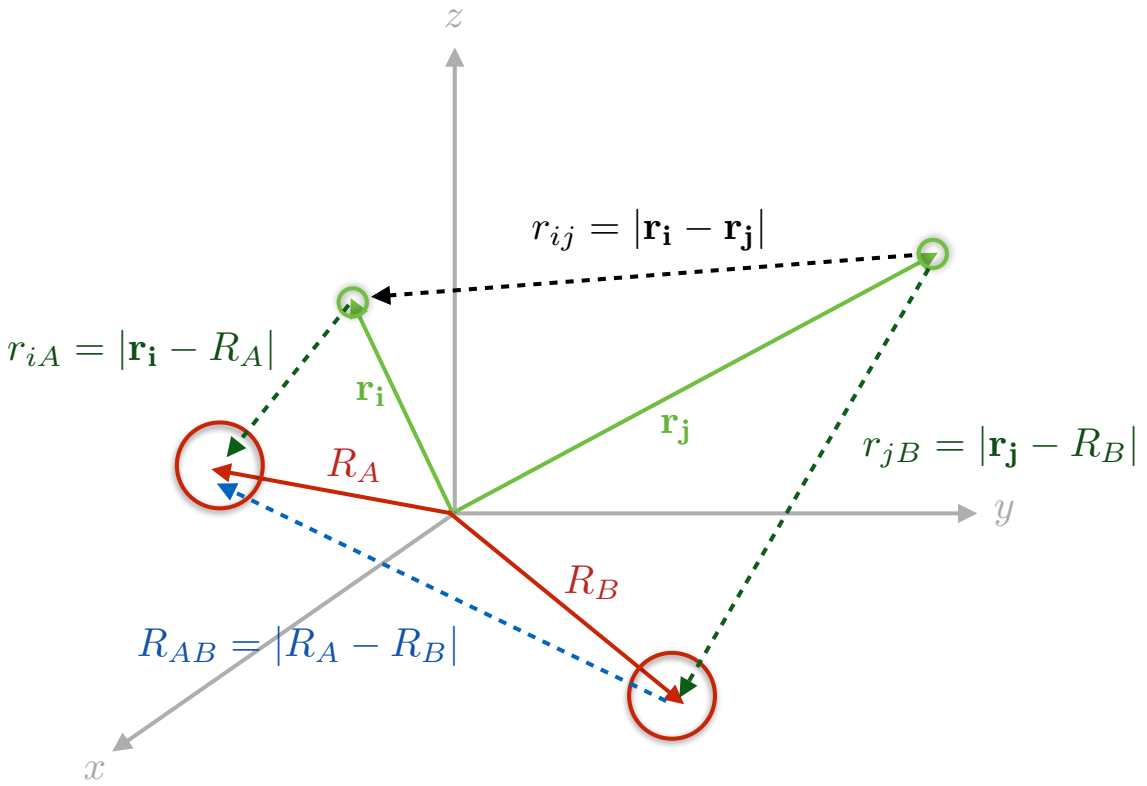


Figure 1.1: Depiction of each of the relevant interactions from the Hamiltonian operator

interpretation, and state that the electronic and nuclear motions are separable and thus, the wavefunction can be written as

$$\Psi(\boldsymbol{r_i}, \boldsymbol{R_A}) = \Psi^{\text{el}}(\boldsymbol{r_i}; \boldsymbol{R_A})\Psi^{\text{nuc}}(\boldsymbol{R_A}) \tag{1.9}$$

For most chemical applications, we can ignore the nuclear wavefunction, $\Psi^{\text{nuc}}(\boldsymbol{R_A})$, and simply concern ourselves with the electronic wavefunction, $\Psi^{\text{el}}(\boldsymbol{r_i}; \boldsymbol{R_A})$. However, notice there is a semicolon in between the $\boldsymbol{r_i}$ and the $\boldsymbol{R_A}$. This form designates everything before the semicolon as a variable of the function, and anything after it to be a parametric dependence. By envoking the Born-Oppenheimer approxima-

tion, the wavefunction now only deals with the positions of the fixed nuclei, and thus, these vectors are now parameters of the electronic wavefunction.

Returning to equation (1.8), there are simplifications that can be made. Since the nuclei have fixed positions, their kinetic energy would be null, and thus the second term in the Hamiltonian operator vanishes. Furthermore, the nuclear repulsion term (fifth and final term) is now comprised solely of constants. A constant operator does not affect the eigenvectors of an eigenvector/eigenvalue problem. It simply adjusts the eigenvalues by that constant, and thus, it can be dealt with after the eigenvectors and eigenvalues are determined. By making these two adjustments, we can simplify the Hamiltonian operator to its electronic counterpart, $\hat{H}^{\text{el}}$, as

$$\hat{H}^{\text{el}} \equiv -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \tag{1.10}$$

By solving the electronic Schrödinger equation

$$\hat{H}^{\text{el}}\Psi^{\text{el}}(\boldsymbol{r_i};\boldsymbol{R_A}) = E^{\text{el}}\Psi^{\text{el}}(\boldsymbol{r_i};\boldsymbol{R_A}) \tag{1.11}$$

the electronic wavefunction, $\Psi^{\text{el}}(\boldsymbol{r_i};\boldsymbol{R_A})$ and electronic energy, E^{el}, are determined. To obtain the total molecular energy, however, the nuclear repulsion energy must be considered. Thus, following the solution to the electronic Schrödinger equation, E^{el} must be combined with the nuclear repulsion component as follows:[6]

$$E = E^{\text{el}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{1.12}$$

As we will only be concerned with the electronic versions of the wavefunction and Hamiltonian operator, from this point forward we will drop the '*el*' designation and it can be assumed that all references to the wavefunction or Hamiltonian operator are referring to the electronic components, unless otherwise stated.

1.3 Solutions to the Schrödinger Equation

By applying the Born-Oppenheimer approximation, the Schrödinger equation was greatly simplified by removing the nuclear position variables. However, solving the equation remains a daunting task, one that is unachievable in principle for nearly all chemical systems. While the solution of the hydrogen atom has been known for many years, its solution would still involve several pages of derivations and requires a familiarity with advanced mathematical functions. Nonetheless, this is the only true chemical system for which there is an exact, closed form solution to the Schrödinger equation. Thus, much of the early work in computational chemistry was focused on the development of methods for obtaining approximate solutions to the Schrödinger equation. Even today, many theorists focus on the development of novel theoretical models to obtain more accurate solutions to the Schrödinger equation. The next few sections will highlight those methods that were fundamental in the development of more accurate models as well as those which are applied in the work presented in this thesis.

1.4 Variational Theorem

In a situation where the exact answer is unachievable, approximations must be made. But without knowing the true answer, how can one assess the accuracy of said approximation? For many methods in computational chemistry, the variational theorem is the tool used for this assessment. The variational theorem states that for any normalized approximate wavefunction, Ψ_G, that satisfies all boundary criteria of the problem, the total energy of that wavefunction, E_G, will always be greater than or equal to the exact ground state energy of that system, E_0.[11] Mathematically, this is written as

$$\langle \Psi_G | \hat{H} | \Psi_G \rangle = E_G \geq E_0 \tag{1.13}$$

Thus, for any method to which the variational theorem applies, it can be certain that out of a series of trial wavefunctions, Ψ_G, the one that yields the lowest energy is that which is a better approximation to the true wavefunction. Furthermore, one can be certain that the energy of any variational method will never be lower than the true energy. Thus, the error in energy for a variational method will always be positive.

1.5 Hartree Product

The Hartree self-consistent field method, or more simply, the Hartree method was proposed by Douglas Hartree in the 1920s.[13,14] While very rarely mentioned outside of introductory theoretical chemistry courses, it is the predecessor of the far more

commonly utilized process, the Hartree Fock (HF) method.

The Hartree method utilizes the Hartree Product wavefunction, Ψ^{HP}. This method assumes that we can approximate the complete wavefunction as a product of single particle functions, or orbitals. This can be expressed as

$$\Psi^{\text{HP}}(\boldsymbol{x_1}, \ldots, \boldsymbol{x_N}) = \prod_{i=1}^{N} \chi_i(\boldsymbol{x_i}) \tag{1.14}$$

where χ_i denotes a spin orbital for electron i with its combined position-spin coordinate vector, $\boldsymbol{x_i} \equiv (\boldsymbol{r}_i, \omega_i)$ (where ω_i is the spin coordinate). The sum goes over N orbitals which is the total number of electrons in the system. One is commonly taught that orbitals describe up to two electrons. That type of statement, however, is referring to spatial orbitals, ψ. A spin orbital describes a single electron with either spin up, $\alpha(\omega)$, or spin down, $\beta(\omega)$. The following equations demonstrate how the first four spin orbitals can be related to spatial orbitals. This trend continues for all other N spin orbitals.

$$\chi_1(\boldsymbol{x}) = \psi_1(\boldsymbol{r})\alpha(\omega) = \psi_1^{\alpha}(\boldsymbol{r}) \tag{1.15}$$

$$\chi_2(\boldsymbol{x}) = \psi_1(\boldsymbol{r})\beta(\omega) = \psi_1^{\beta}(\boldsymbol{r}) \tag{1.16}$$

$$\chi_3(\boldsymbol{x}) = \psi_2(\boldsymbol{r})\alpha(\omega) = \psi_2^{\alpha}(\boldsymbol{r}) \tag{1.17}$$

$$\chi_4(\boldsymbol{x}) = \psi_2(\boldsymbol{r})\beta(\omega) = \psi_2^{\beta}(\boldsymbol{r}) \tag{1.18}$$

While spin functions have no mathematical form, there are a few properties that should be identified. First and foremost, the spin functions are normalized and

orthogonal to each other. This is expressed mathematically as

$$\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1 \qquad \text{Normalization} \tag{1.19}$$

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle = 0 \qquad \text{Orthogonality} \tag{1.20}$$

As the spin components are typically integrated out for most purposes, these identities greatly simplify expressions as they will nullify many components or make the spin components equal to unity.

There are two major problems with the Hartree method. Neither are obvious from a classical perspective; however, in quantum mechanics, there are two important properties that are being violated here. The first is that the wavefunction of fermionic particles must be antisymmetric.[6,8,11] More simply, if the positions of two electrons are swapped, the resulting wavefunction must be the negative of the first. In other words

$$\Psi(\boldsymbol{x_2}, \boldsymbol{x_1}, \ldots, \boldsymbol{x_N}) = -\Psi(\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_N}) \tag{1.21}$$

There is nothing that guarantees that this antisymmetry principle holds true with the Hartree-Product wavefunction. The second principle of quantum mechanics that is being violated here is that electrons are indistinguishable. While one can easily label and distinguish classical objects, this is simply not possible for electrons. The Hartree-Product wavefunction specifies in which orbital each electron resides which violates the need for indistinguishable particles.

1.6 Hartree-Fock Theory

Following the development of the Hartree method, Fock sought to modify the method to satisfy these two quantum mechanical properties.[15] His solution, while still utilizing single-particle orbitals, was to consider every possible permutation of electrons within each of those orbitals. Mathematically, the simplest way to do that was by using a determinant. These so-called Slater determinants[16,17] are used to construct Hartree-Fock (HF) wavefunctions, Ψ^{HF}, for any N-electron system. The following equation demonstrates the expression of an HF wavefunction for a 2-electron system.

$$\Psi^{\text{HF}}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{x_1}) & \chi_2(\boldsymbol{x_1}) \\ \chi_1(\boldsymbol{x_2}) & \chi_2(\boldsymbol{x_2}) \end{vmatrix} = \frac{1}{\sqrt{2!}} \begin{vmatrix} \chi_1(\boldsymbol{x_1}) & \chi_2(\boldsymbol{x_1}) \\ \chi_1(\boldsymbol{x_2}) & \chi_2(\boldsymbol{x_2}) \end{vmatrix} \tag{1.22}$$

where the $(N!)^{-1/2}$ is a normalization constant dependent on the number of electrons in the system. In this determinant, each column corresponds to a specific spin orbital while each row corresponds to a specific electron. By working out the determinant, momentarily ignoring the normalization constant, we get

$$\begin{vmatrix} \chi_1(\boldsymbol{x_1}) & \chi_2(\boldsymbol{x_1}) \\ \chi_1(\boldsymbol{x_2}) & \chi_2(\boldsymbol{x_2}) \end{vmatrix} = \chi_1(\boldsymbol{x_1})\chi_2(\boldsymbol{x_2}) - \chi_2(\boldsymbol{x_1})\chi_1(\boldsymbol{x_2}) \tag{1.23}$$

From the right hand side of this expression, one can note that every electron is now occupying every orbital. Thus, Ψ^{HF} satisfies the idea that electrons must be

indistinguishable. It remains that this wavefunction must be antisymmetric. This is confirmed as follows

$$\chi_1(\boldsymbol{x_2})\chi_2(\boldsymbol{x_1}) - \chi_2(\boldsymbol{x_2})\chi_1(\boldsymbol{x_1}) = -(\chi_1(\boldsymbol{x_1})\chi_2(\boldsymbol{x_2}) - \chi_2(\boldsymbol{x_1})\chi_1(\boldsymbol{x_2})) \tag{1.24}$$

and thus

$$\Psi^{\text{HF}}(\boldsymbol{x_2}, \boldsymbol{x_1}) = -\Psi^{\text{HF}}(\boldsymbol{x_1}, \boldsymbol{x_2}) \tag{1.25}$$

Expressing the wavefunction as a Slater determinant has a number of useful properties. First and foremost, it addresses the issues in the Hartree-Product wavefunction.[13,14] However, consider a few important properties of determinants. If two columns of a determinant are identical, the determinant is equal to zero. In this context, two equivalent columns corresponds to two identical orbitals which violates the Pauli Exclusion principle. Secondly, by interchanging two columns in a determinant, the resulting determinant is equivalent but multiplied by a factor of -1. This property ensures that the wavefunction will be antisymmetric.

While these two properties were demonstrated using a wavefunction for a two electron system, they hold true for any general system. The general N-electron

Hartree-Fock wavefunction is given by

$$\Psi^{\text{HF}}(\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{x_1}) & \chi_2(\boldsymbol{x_1}) & \cdots & \chi_N(\boldsymbol{x_1}) \\ \chi_1(\boldsymbol{x_2}) & \chi_2(\boldsymbol{x_2}) & \cdots & \chi_N(\boldsymbol{x_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\boldsymbol{x_N}) & \chi_2(\boldsymbol{x_N}) & \cdots & \chi_N(\boldsymbol{x_N}) \end{vmatrix} \tag{1.26}$$

While the HF method does address the two previously described issues with the Hartree method, there remains an issue that has yet to be discussed. This issue is known as electron correlation.[6] There are two types of correlation: Fermi correlation and Coulomb correlation. The HF method addresses the issue of Fermi correlation which is the concept that two electrons with the same spin will never be at the same point in space at the same time. This type of correlation is absent in the Hartree method, but is addressed exactly in HF theory by preventing two spin orbitals from being identical (determinant would equal zero).

However, the HF method does not account for Coulomb correlation. This form of correlation energy is the result of electrons interacting and avoiding one another.[6] Hartree-Fock theory applies a mean field approach, where each individual electron does not *see* each other electron, but simply feels the mean electric field generated by all other electrons in the system. While Coulomb correlation only accounts for approximately 1% of total electronic energies, it is on the same order of magnitude as many reaction energies. Thus, to study reaction energetics using Hartree-Fock theory and expecting great results would be like buying a lottery ticket and expect-

ing to win. It is possible; however, it is not overly probable.

The full derivation of the Hartree-Fock equations is outside of the scope of this thesis. The interested reader is directed to the works of Szabo and Ostlund[11] as well as Levine.[6] However, the basics of the method will be explained herein with the example of a two-electron system. The equations will then be generalized to the N-electron case.

Expectation values were briefly introduced, although not referenced by name, in the discussion of the Variational Theorem and equation (1.13). An expectation value or average value of an operator, $\hat{O}$, is given by

$$\frac{\langle\Psi|\hat{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int \Psi^*(\boldsymbol{x})\hat{O}\,\Psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}}{\int \Psi^*(\boldsymbol{x})\,\Psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x}} = \langle\hat{O}\rangle \tag{1.27}$$

However, in the event that the wavefunction is normalized, the denominators are equal to unity, and the equation simplifies to

$$\langle\Psi|\hat{O}|\Psi\rangle = \int \Psi^*(\boldsymbol{x})\hat{O}\,\Psi(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \langle\hat{O}\rangle \tag{1.28}$$

In the event that the operator is the Hamiltonian, the expectation value or $\langle\hat{H}\rangle$ would be equal to the ground state energy of the chemical system. Applying this theory to the Hartree-Fock problem yields

$$E^{\mathrm{HF}} = \langle\Psi^{\mathrm{HF}}|\hat{H}|\Psi^{\mathrm{HF}}\rangle \tag{1.29}$$

The problem can be simplified by separating the Hamiltonian into two separate

pieces. After applying the Born-Oppenheimer approximation, recall that we are left with

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{1.30}$$

The components in green are the one-electron operators and can be labelled $\hat{H}_1$ which is defined as

$$\hat{H}_1 \equiv \sum_{i=1}^{N}\hat{H}_i^{\text{core}} \tag{1.31}$$

where the core Hamiltonian, $\hat{H}^{\text{core}}$, is given by

$$\hat{H}_i^{\text{core}} \equiv -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \tag{1.32}$$

The two-electron Hamiltonian, $\hat{H}_2$ is then simply

$$\hat{H}_2 \equiv \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{1.33}$$

Using these new definitions, the Hartree Fock energy defined in equation (1.29) can be rewritten as

$$E^{\text{HF}} = \langle\Psi^{\text{HF}}|\hat{H}_1|\Psi^{\text{HF}}\rangle + \langle\Psi^{\text{HF}}|\hat{H}_2|\Psi^{\text{HF}}\rangle \tag{1.34}$$

$$E^{\text{HF}} = \langle\Psi^{\text{HF}}|\sum_{i=1}^{N}\hat{H}_i^{\text{core}}|\Psi^{\text{HF}}\rangle + \langle\Psi^{\text{HF}}|\hat{H}_2|\Psi^{\text{HF}}\rangle \tag{1.35}$$

Considering that electrons are indistinguishable particles,[6] the energy of a single electron should be $1/N$ times of the total one-electron energy. Using this logic,

(1.35) is simplified to

$$E^{\mathrm{HF}} = N\langle\Psi^{\mathrm{HF}}|\hat{H}_1^{\mathrm{core}}|\Psi^{\mathrm{HF}}\rangle + \langle\Psi^{\mathrm{HF}}|\hat{H}_2|\Psi^{\mathrm{HF}}\rangle \tag{1.36}$$

To solve for the first term in this expression, we can substitute the right hand side of equation (1.23) into (1.32) to obtain

$$\begin{aligned}\langle\Psi^{\mathrm{HF}}|\hat{H}_1^{\mathrm{core}}|\Psi^{\mathrm{HF}}\rangle =& \frac{1}{N!}\int\int[\chi_1^*(\boldsymbol{x_1})\chi_2^*(\boldsymbol{x_2}) - \chi_2^*(\boldsymbol{x_1})\chi_1^*(\boldsymbol{x_2})] \\ & \hat{H}_1^{\mathrm{core}}[\chi_1(\boldsymbol{x_1})\chi_2(\boldsymbol{x_2}) - \chi_2(\boldsymbol{x_1})\chi_1(\boldsymbol{x_2})]\mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2}\end{aligned} \tag{1.37}$$

where we have reintroduced the normalization constant. Expanding this expression by multiplying everything out will yield four separate integrals as follows

$$\begin{aligned}\langle\Psi^{\mathrm{HF}}|\hat{H}_1^{\mathrm{core}}|\Psi^{\mathrm{HF}}\rangle = & \frac{1}{2}\int\int[\chi_1^*(\boldsymbol{x_1})\chi_2^*(\boldsymbol{x_2})\hat{H}_1^{\mathrm{core}}\chi_1(\boldsymbol{x_1})\chi_2(\boldsymbol{x_2})]\mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2} \\ & - \frac{1}{2}\int\int[\chi_1^*(\boldsymbol{x_1})\chi_2^*(\boldsymbol{x_2})\hat{H}_1^{\mathrm{core}}\chi_2(\boldsymbol{x_1})\chi_1(\boldsymbol{x_2})]\mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2} \\ & - \frac{1}{2}\int\int[\chi_2^*(\boldsymbol{x_1})\chi_1^*(\boldsymbol{x_2})\hat{H}_1^{\mathrm{core}}\chi_1(\boldsymbol{x_1})\chi_2(\boldsymbol{x_2})]\mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2} \\ & + \frac{1}{2}\int\int[\chi_2^*(\boldsymbol{x_1})\chi_1^*(\boldsymbol{x_2})\hat{H}_1^{\mathrm{core}}\chi_2(\boldsymbol{x_1})\chi_1(\boldsymbol{x_2})]\mathrm{d}\boldsymbol{x_1}\mathrm{d}\boldsymbol{x_2}\end{aligned} \tag{1.38}$$

As $\hat{H}_1^{\mathrm{core}}$ only contains terms that operate on electron one, any terms involving $\boldsymbol{x_2}$ can be factored out and integrated separately. For the sake of simplicity, we will now replace $\boldsymbol{x_i}$ with $\boldsymbol{i}$. This will yield the following four terms based on the

previous expression:

$$\frac{1}{2}\int \chi_1^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_1(\mathbf{1})\mathrm{d}\mathbf{1}\int \chi_2^*(\mathbf{2})\chi_2(\mathbf{2})\mathrm{d}\mathbf{2} = \frac{1}{2}\int \chi_1^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_1(\mathbf{1})\mathrm{d}\mathbf{1} \tag{1.39}$$

$$-\frac{1}{2}\int \chi_1^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_2(\mathbf{1})\mathrm{d}\mathbf{1}\int \chi_2^*(\mathbf{2})\chi_1(\mathbf{2})\mathrm{d}\mathbf{2} = 0 \tag{1.40}$$

$$-\frac{1}{2}\int \chi_2^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_1(\mathbf{1})\mathrm{d}\mathbf{1}\int \chi_1^*(\mathbf{2})\chi_2(\mathbf{2})\mathrm{d}\mathbf{2} = 0 \tag{1.41}$$

$$\frac{1}{2}\int \chi_2^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_2(\mathbf{1})\mathrm{d}\mathbf{1}\int \chi_1^*(\mathbf{2})\chi_1(\mathbf{2})\mathrm{d}\mathbf{2} = \frac{1}{2}\int \chi_2^*(\mathbf{1})\hat{H}_1^{\text{core}}\chi_2(\mathbf{1})\mathrm{d}\mathbf{1} \tag{1.42}$$

The orthonormality of the spin functions,[11] $\alpha(\omega)$ and $\beta(\omega)$, and the spatial orbitals themselves were utilized to simplify these four terms. In the first and fourth terms, the same spin function is involved in the integral over electron 2 and thus, the integral equates to unity due to the normalization. For the second and third terms, however, the opposite spin functions are involved in the electron 2 integral which causes the integral the vanish and the full term to equal zero.

The terms that were not nullified through integration can be simplified further by separating the spin orbitals, $\chi(\boldsymbol{x})$, to their spatial orbital, $\psi(\boldsymbol{r})$, and spin function components. By doing so, the following expressions are obtained

$$\begin{aligned}
&\frac{1}{2}\int\int \psi_1^*(\boldsymbol{r_1})\alpha^*(\omega_1)\hat{H}_1^{\text{core}}\psi_1(\boldsymbol{r_1})\alpha(\omega_1)\mathrm{d}\boldsymbol{r_1}\mathrm{d}\omega_1 = \\
&\frac{1}{2}\int \psi_1^*(\boldsymbol{r_1})\hat{H}_1^{\text{core}}\psi_1(\boldsymbol{r_1})\mathrm{d}\boldsymbol{r_1}\int \alpha^*(\omega_1)\alpha(\omega_1)\mathrm{d}\omega_1 = \\
&\frac{1}{2}\int \psi_1^*(\boldsymbol{r_1})\hat{H}_1^{\text{core}}\psi_1(\boldsymbol{r_1})\mathrm{d}\boldsymbol{r_1} = H_{11}
\end{aligned} \tag{1.43}$$

An analogous procedure can be done for term 4 which would yield another H_{11}

term. It should be noted that when working with spatial orbitals, said term would be considered the H_{22} term; however, recall that there are two spin orbitals per spatial orbital. Thus, We can now succinctly express the expectation value of the one-electron Hamiltonian (i.e. the one-electron component to the energy) as

$$\langle \Psi^{\text{HF}} | \hat{H}_1 | \Psi^{\text{HF}} \rangle = \int \psi_i^*(\boldsymbol{r_1}) \hat{H}_1^{\text{core}} \psi_i(\boldsymbol{r_1}) \mathrm{d}\boldsymbol{r_1} = 2 \sum_{i=1}^{N/2} H_{ii} \tag{1.44}$$

Based on this result, the closed-shell HF energy can now be expressed as

$$E^{\text{HF}} = 2 \sum_{i=1}^{N/2} H_{ii} + \langle \Psi^{\text{HF}} | \hat{H}_2 | \Psi^{\text{HF}} \rangle \tag{1.45}$$

Thus, it remains to determine how to calculate the two-electron component of the energy. As previously noted, the full derivation of the HF equation will not be conducted here. The one-electron component was fully derived as it is quite simple; however, only the highlights of the two electron component will be shown.

Recall that the two-electron component to the HF energy is given by

$$\langle \Psi^{\text{HF}} | \hat{H}_2 | \Psi^{\text{HF}} \rangle = \langle \Psi^{\text{HF}} | \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} | \Psi^{\text{HF}} \rangle \tag{1.46}$$

After substituting the full form of the HF wavefunction, and integrating out the spin components of the spatial orbitals, only two types of integrals remain. These are known as the Coulomb integrals, J_{ij}, and exchange integrals K_{ij}. The Coulomb

integrals are given by

$$J_{ij} = \int\int \psi_i^*(\mathbf{1})\psi_j^*(\mathbf{2})\frac{1}{r_{12}}\psi_i(\mathbf{1})\psi_j(\mathbf{2})\mathrm{d}\mathbf{1}\mathrm{d}\mathbf{2}$$
$$J_{ij} = \int\int |\psi_i(\mathbf{1})|^2\frac{1}{r_{12}}|\psi_j(\mathbf{2})|^2\mathrm{d}\mathbf{1}\mathrm{d}\mathbf{2} \tag{1.47}$$

As $|\psi_i|^2$ describes the electron density of a molecular orbital i, J_{ij} represent the interaction between the electron density of two separate molecular orbitals. Thus, one can interpret the Coulomb integrals as the Coulombic repulsion between electrons, hence the name. The other type of integral remaining are exchange integrals, which are calculated by

$$K_{ij} = \int\int \psi_i^*(\mathbf{1})\psi_j^*(\mathbf{2})\frac{1}{r_{12}}\psi_j(\mathbf{1})\psi_i(\mathbf{2})\mathrm{d}\mathbf{1}\mathrm{d}\mathbf{2} \tag{1.48}$$

Comparing the terms on the left of the operator and those to the right of the operator, one can notice that the only difference is that the electrons are switched or *exchanged* which provides an explanation for the name. The presence of these integrals is due to the antisymmetry of the wavefunction. While the Coulomb integrals can be interpreted as the Coulombic repulsion between electrons, there is no physical interpretation for the exchange integrals.[6,11] They are a quantum mechanical effect with no classical interpretation, but are an important part of the two-electron energy nonetheless.

Exchange is a phenomenon that only occurs between same spin electrons,[6] where-as Coulombic interactions occur between every single pair of electrons. In

terms of these two components, the two-electron energy can be described by

$$\langle \Psi^{\text{HF}} | \hat{H}_2 | \Psi^{\text{HF}} \rangle = \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2J_{ij} - K_{ij} \tag{1.49}$$

Combining equations (1.44) and (1.49) yields the full expression for the closed-shell Hartree-Fock energy

$$E^{\text{HF}} = 2\sum_{i=1}^{N/2} H_{ii} + \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2J_{ij} - K_{ij} \tag{1.50}$$

The only issue with this expression is that the identity of the molecular orbitals, ψ, are unknown. Solutions to this were proposed separately by Roothaan[18] and Hall.[19] By approximating each molecular orbital as a linear combination of atomic orbitals

$$\psi_i(\boldsymbol{r}) = \sum_{\mu=1}^{K} c_{\mu,i}\phi_\mu(\boldsymbol{r}) \tag{1.51}$$

where $c_{\mu,i}$ is the amount of each of the K basis functions, $\phi_\mu(\boldsymbol{r})$, that is used to approximate each molecular orbital, $\psi_i(\boldsymbol{r})$. The functional form of these basis functions will be explored in a later section. For now, we will use these functions to rewrite the HF energy in a usable form.

Substituting (1.51) into (1.50), eventually yields the final expression for the HF energy, E^{HF}:

$$E^{\text{HF}} = \sum_{\mu=1}^{K}\sum_{\nu=1}^{K} P_{\mu\nu}H_{\mu\nu}^{\text{core}} + \frac{1}{2}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} P_{\mu\nu}P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\sigma|\lambda\nu)] \tag{1.52}$$

But how do we get to this final expression? Let's first consider the 1-electron component to the energy:

$$\langle \Psi^{\text{HF}} | \hat{H}_1 | \Psi^{\text{HF}} \rangle = 2 \sum_{i=1}^{N/2} H_{ii} = 2 \sum_{i=1}^{N/2} \langle \psi_i | \hat{H}_{\text{core}} | \psi_i \rangle \tag{1.53}$$

Using the LCAO representation for the molecular orbitals, equation (1.51) can be rewritten as:

$$2 \sum_{i=1}^{N/2} \langle \psi_i | \hat{H}_{\text{core}} | \psi_i \rangle = 2 \sum_{i=1}^{N/2} \langle \sum_{\mu=1}^{K} c_{\mu,i} \phi_\mu | \hat{H}_{\text{core}} | \sum_{\nu=1}^{K} c_{\nu,i} \phi_\nu \rangle \tag{1.54}$$

Since the coefficients are not affected by the core Hamiltonian operator, they can be brought outside of the inner product expression as such:

$$2 \sum_{i=1}^{N/2} \langle \psi_i | \hat{H}_{\text{core}} | \psi_i \rangle = 2 \sum_{i=1}^{N/2} \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} c^*_{\mu,i} c_{\nu,i} \langle \phi_\mu | \hat{H}_{\text{core}} | \phi_\nu \rangle \tag{1.55}$$

To simply this expression further, we introduce the density matrix,[11] $P_{\mu\nu}$, which is defined as

$$P_{\mu\nu} = 2 \sum_{i}^{N/2} c^*_{\mu,i} c_{\nu,i} \tag{1.56}$$

Applying this definition of $P_{\mu\nu}$ to equation (1.55) gives

$$2 \sum_{i=1}^{N/2} \langle \psi_i | \hat{H}_{\text{core}} | \psi_i \rangle = \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} P_{\mu\nu} \langle \phi_\mu | \hat{H}_{\text{core}} | \phi_\nu \rangle \tag{1.57}$$

Now we have an expression for the 1-electron HF energy in terms of atomic orbitals,

ϕ_μ, and their respective coefficients, $c_{\mu,i}$. The form of these atomic orbitals and how these coefficients are obtained will be discussed in the coming sections. For now, we turn our attention to the 2-electron component of the energy.

$$\langle \Psi^{\text{HF}} | \hat{H}_2 | \Psi^{\text{HF}} \rangle = \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2J_{ij} - K_{ij} \tag{1.58}$$

Let us first consider the Coulombic portion (J_{ij}) of the two-electron energy. By once again applying the LCAO expansion to the Coulombic portion this time, we obtain

$$\sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2J_{ij} = \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2\langle \psi_i(1)\psi_i(1) | \frac{1}{r_{12}} | \psi_j(2)\psi_j(2) \rangle \tag{1.59}$$

$$= \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2\langle \sum_{\mu=1}^{K} c_{\mu,i}\phi_\mu(1) \sum_{\nu=1}^{K} c_{\nu,i}\phi_\nu(1) | \frac{1}{r_{12}} | \sum_{\lambda=1}^{K} c_{\lambda,j}\phi_\lambda(2) \sum_{\sigma=1}^{K} c_{\sigma,j}\phi_\sigma(2) \rangle \tag{1.60}$$

This can be simplified, as before, by bringing the coefficients outside of the inner product:

$$\sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2J_{ij} = \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} 2 \sum_{\mu=1}^{K} \sum_{\nu=1}^{K} c_{\mu,i}c_{\nu,i} \sum_{\lambda=1}^{K} \sum_{\sigma=1}^{K} c_{\lambda,j}c_{\sigma,j} \langle \phi_\mu(1)\phi_\nu(1) | \frac{1}{r_{12}} | \phi_\lambda(2)\phi_\sigma(2) \rangle \tag{1.61}$$

Using the definition of the density matrix given in equation (1.56), this simplifies to

$$\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} 2J_{ij} = 2\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} \tfrac{1}{2}P_{\mu\nu}\tfrac{1}{2}P_{\lambda\sigma}\langle\phi_\mu(1)\phi_\nu(1)|\frac{1}{r_{12}}|\phi_\lambda(2)\phi_\sigma(2)\rangle \tag{1.62}$$

Often times, chemists will use a shorthand notation to simplify the expression for the two integrals.[11] In this shorthand, the inner product is shown as:

$$\langle\phi_\mu(1)\phi_\nu(1)|\frac{1}{r_{12}}|\phi_\lambda(2)\phi_\sigma(2)\rangle = (\mu\nu|\lambda\sigma) \tag{1.63}$$

Applying this shorthand to equation (1.62) yields a compact expression for the Coulombic portion of the HF energy

$$\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} 2J_{ij} = \frac{1}{2}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} P_{\mu\nu}P_{\lambda\sigma}(\mu\nu|\lambda\sigma) \tag{1.64}$$

The only remaining portion of the HF energy is the exchange component (K_{ij}). Recall that the exchange energy is given by

$$-\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} K_{ij} = -\sum_{i=1}^{N/2}\sum_{j=1}^{N/2}\langle\psi_i(1)\psi_j(2)|\frac{1}{r_{12}}|\psi_j(1)\psi_i(2)\rangle \tag{1.65}$$

Replacing the molecular orbitals with the LCAO expansion one last time gives

$$-\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} K_{ij} = -\sum_{i=1}^{N/2}\sum_{j=1}^{N/2}\langle\sum_{\mu=1}^{K} c_{\mu,i}\phi_\mu(1)\sum_{\lambda=1}^{K} c_{\lambda,j}\phi_\lambda(2)|\frac{1}{r_{12}}|\sum_{\sigma=1}^{K} c_{\sigma,j}\phi_\sigma(1)\sum_{\nu=1}^{K} c_{\nu,i}\phi_\nu(2)\rangle \tag{1.66}$$

By bringing all of the coefficients outside of the inner product and applying the definition of the density matrix, we obtain

$$-\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} K_{ij} = -\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} \tfrac{1}{2}P_{\mu\nu}\tfrac{1}{2}P_{\lambda\sigma}\langle\phi_\mu(1)\phi_\lambda(2)|\frac{1}{r_{12}}|\phi_\sigma(1)\phi_\nu(2)\rangle \tag{1.67}$$

or with the shorthand notation defined in equation (1.63)

$$-\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} K_{ij} = -\frac{1}{4}\sum_{\mu=1}^{K}\sum_{\nu=1}^{K}\sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} P_{\mu\nu}P_{\lambda\sigma}(\mu\sigma|\lambda\nu) \tag{1.68}$$

Combining all of the components of the HF energy expressed in equations (1.57), (1.64), and (1.68) yields the final expression for the Hartree-Fock energy

$$E^{\text{HF}} = \sum_{\mu\nu}^{K} P_{\mu\nu}\langle\phi_\mu|\hat{H}_{\text{core}}|\phi_\nu\rangle + \frac{1}{2}\sum_{\mu\nu\lambda\sigma}^{K} P_{\mu\nu}P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\sigma|\lambda\nu)] \tag{1.69}$$

where the summations are combined into one for the sake of simplicity. This is identical to equation (1.52) which was shown at the start of this discussion regarding the introduction of the atomic orbital basis set.

1.7 Hartree-Fock Equations

While equation (1.69) describes how to obtain the Hartree-Fock energy, it does not tell us how to determine the atomic orbital coefficients, $c_{\mu,i}$, for the LCAO expansions. These are determined through the Hartree-Fock equations. The complete derivation of the HF equations will not be included herein, but for the interested reader, a thorough derivation is provided by Szabo and Ostlund.[11]

The basis for determining the atomic orbital coefficients is taking the first derivative of the energy with respect to the coefficients, $c_{\mu,i}$ and setting it equal to zero. As with any function, this will determine the minimum of the function (or in some cases the maximum). The Hartree-Fock equations are the result of performing this operation on the energy expression while requiring that the set of orbitals remain orthonormal, i.e. orthogonal to one another and normalized. By representing our molecular orbitals in the HF equations using the LCAO approach, we obtain the Roothaan-Hall equations[18,19] which is commonly expressed as

$$\boldsymbol{FC} = \boldsymbol{SCE} \tag{1.70}$$

where $\boldsymbol{F}$, $\boldsymbol{C}$, $\boldsymbol{S}$, and $\boldsymbol{E}$ are the Fock matrix, the coefficient matrix, the overlap matrix, and the energy matrix, respectively. The energy matrix, $\boldsymbol{E}$, is a diagonal matrix in which the diagonal elements correspond to the energies of the molecular orbitals, ε_i. The form of each of the remaining matrices will be discussed in the coming paragraphs.

The Fock matrix, $\boldsymbol{F}$, is a $K \times K$ matrix comprised of elements, $F_{\mu\nu}$ which are given by

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + \sum_{\lambda=1}^{K}\sum_{\sigma=1}^{K} P_{\lambda\sigma}[(\mu\nu|\lambda\sigma) - \tfrac{1}{2}(\mu\lambda|\nu\sigma)]] \tag{1.71}$$

where $H_{\mu\nu}^{\text{core}}$ is an element of the core Hamiltonian matrix which are calculated from

$$H_{\mu\nu}^{\text{core}} = \int \phi_\mu^*(\boldsymbol{r}_1)\hat{H}_1^{\text{core}}\phi_\nu(\boldsymbol{r}_1)\mathrm{d}\boldsymbol{r}_1 \tag{1.72}$$

In this equation, $\hat{H}_1^{\text{core}}$ is the core Hamiltonian operator that was defined in equation (1.32).

The overlap matrix, $\boldsymbol{S}$, is another $K \times K$ matrix where the elements $S_{\mu\mu}$ (i.e. the diagonal elements) are equal to unity. As the name suggests, it measures the level of overlap between two *basis functions*. The emphasis on basis functions is important, as were these molecular orbitals, $\boldsymbol{S}$ would be equal to the identity matrix due to the requirement that the MOs be an orthonormal set. The elements of the overlap matrix are determined from[11]

$$S_{\mu\nu} = \int \phi_\mu^*(\boldsymbol{r}_1)\phi_\mu(\boldsymbol{r}_1)\mathrm{d}\boldsymbol{r}_1 \tag{1.73}$$

Finally, the coefficient matrix, $\boldsymbol{C}$, details the contribution of each atomic orbital or basis function to the molecular orbitals of the system. In this matrix, the rows describe the basis functions, while the columns represent the molecular orbitals. For instance, in the following matrix for a chemical system with K basis functions

$$\boldsymbol{C} = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,K} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ c_{K,1} & c_{K,2} & \cdots & c_{K,K} \end{pmatrix} \tag{1.74}$$

$c_{2,1}$ is the coefficient describing how much of basis function 2, ϕ_2, contributes to molecular orbital 1, ψ_1, while $c_{1,2}$ would describe how much of basis function 2, ϕ_1, contributes to molecular orbital 2, ψ_2.

While the Roothaan-Hall equations appear to be a simple eigenvalue problem, the Fock matrix is dependent on its own eigenfunctions, and thus, this problem must be solved iteratively. This process is known as a self-consistent field model as the process is continued iteratively until the the orbital energies converge, or become self-consistent.[11]

In the self-consistent field method, the first step involves the transformation of the overlap matrix, $\boldsymbol{S}$, into the identity matrix, $\mathbb{1}$. There are a few different ways to achieve this which are described by Szabo and Ostlund,[11] but the approach that will be covered here is that of symmetrical orthogonalization. This involves taking the inverse square root of the overlap matrix, denoted by $\boldsymbol{S}^{-1/2}$, but is often represented as $\boldsymbol{X}$. $\boldsymbol{X}$ is defined such that

$$\boldsymbol{X}^{\dagger}\boldsymbol{S}\boldsymbol{X} = \mathbb{1} \tag{1.75}$$

where $\boldsymbol{X}^{\dagger}$ is the conjugate transpose of $\boldsymbol{X}$. Using this new matrix, $\boldsymbol{X}^{\dagger}$, it is multiplied on the left of the Roothaan-hall equations to give

$$\boldsymbol{X}^{\dagger}\boldsymbol{F}\boldsymbol{C} = \boldsymbol{X}^{\dagger}\boldsymbol{S}\boldsymbol{C}\boldsymbol{E} \tag{1.76}$$

Taking advantage of the fact that a matrix multiplied by its inverse is equal to the identity matrix, as summarized below,

$$\boldsymbol{X}\boldsymbol{X}^{-\mathbf{1}} = \mathbb{1} \tag{1.77}$$

we can insert this anywhere in equation (1.76) without changing either side of the

equation. By multiplying $\boldsymbol{F}$ and $\boldsymbol{S}$ in (1.76)on the right by $\boldsymbol{X}\boldsymbol{X}^{-1}$ gives

$$[\boldsymbol{X}^{\dagger}\boldsymbol{F}\boldsymbol{X}](\boldsymbol{X}^{-1}\boldsymbol{C}) = \boldsymbol{X}^{\dagger}\boldsymbol{S}\boldsymbol{X}(\boldsymbol{X}^{-1}\boldsymbol{C})\boldsymbol{E} \tag{1.78}$$

By applying the property of $\boldsymbol{X}$ described in (1.75), this can be simplified to

$$[\boldsymbol{X}^{\dagger}\boldsymbol{F}\boldsymbol{X}](\boldsymbol{X}^{-1}\boldsymbol{C}) = (\boldsymbol{X}^{-1}\boldsymbol{C})\boldsymbol{E} \tag{1.79}$$

Introducing two new matrices, $\boldsymbol{F'}$ and $\boldsymbol{C'}$, which are defined by the terms in square and round brackets, respectively, this equation simplifies further to

$$\boldsymbol{F'}\boldsymbol{C'} = \boldsymbol{C'}\boldsymbol{E} \tag{1.80}$$

Using this new form of the Roothaan-Hall equations, the new pseudo-coefficient matrix, $\boldsymbol{C'}$ as well as the energy matrix, can be determined through the diagonalization of $\boldsymbol{F}'$. $\boldsymbol{C}'$ and $\boldsymbol{E}$ are then the eigenvectors and eigenvalues, respectively, of the diagonalized $\boldsymbol{F}'$ matrix.

As noted previously, $\boldsymbol{C'}$, is not the true MO coefficient matrix. As $\boldsymbol{C'}$ is determined from

$$\boldsymbol{C}' = \boldsymbol{X}^{-1}\boldsymbol{C} \tag{1.81}$$

The new coefficient matrix, $\boldsymbol{C}$, is then determined by multiplying this equation by $\boldsymbol{X}$ on the left as follows:

$$\boldsymbol{X}\,\boldsymbol{C'} = \boldsymbol{C}_{\text{new}} \tag{1.82}$$

This new set of coefficients can then be used to calculate a new Fock matrix, $\boldsymbol{F}$, which in turn is used to calculate another new set of coefficients. This process is repeated until the new coefficients are identical to the previous set within a specified level of convergence. Thus, in summary, the SCF method consists of the following procedure[11]:

1. Choose an appropriate set of atomic orbitals
2. Calculate all required integrals for $S_{\mu\nu}$, $H_{\mu\nu}^{\text{core}}$, $(\mu\nu|\lambda\sigma)$, and $(\mu\lambda|\sigma\nu)$
3. Determine the inverse square root of the overlap matrix, $\boldsymbol{X}$
4. Obtain a “guess" at the molecular orbital coefficients to determine $\boldsymbol{C}$
5. Use $\boldsymbol{C}$ to calculate $P_{\lambda\sigma}$ using equation (1.56)
6. Obtain the Fock matrix, from equation (1.71)
7. Calculate the transformed Fock matrix, $\boldsymbol{F}'$, from $\boldsymbol{F}' = \boldsymbol{X}^{\dagger}\boldsymbol{F}\boldsymbol{X}$
8. Diagonalize $\boldsymbol{F}'$ to determine $\boldsymbol{C}'$ and $\boldsymbol{E}$
9. Calculate the new coefficient matrix, $\boldsymbol{C}_{\text{new}}$ from equation (1.82)
10. Repeat steps 5-9 until adequate convergence is achieved

The first step of this process requires the choice of an appropriate set of atomic orbitals. These atomic orbitals or basis functions will be discussed in a later section. In the fourth step, a coefficient matrix must be constructed from a “guess". There

are a number of different ways in which to make this guess but they will not be discussed herein. Many different types of guesses are available in the various different quantum chemical software packages.[20–22]

1.8 Unrestricted Hartree-Fock Theory

The discussion regarding Hartree-Fock theory thus far has focused on closed-shell systems (i.e. systems with no unpaired electrons). This is referred to as the restricted HF method (RHF). For systems with unpaired electrons, such as radicals, transition metal complexes, or excited states, this type of analysis does not generally apply.

Earlier in this chapter, we discussed the difference between spin, χ, and spatial, ψ, orbitals. Spin orbitals contain a single electron while spatial orbitals contain two-electrons with differing spin components. The unrestricted HF method (UHF) does not restrict a pair of electrons to be contained within the same spatial orbital. Instead, each spin orbital contains a single electron and are not restricted to be localized in the same region of space. Consider the example of an H_2 molecule.[11] At small internuclear distances, the two electrons in the system are likely interacting with both nuclei and form a covalent bond between the two hydrogen atoms. For this system, the RHF model is completely accurate; however, as the distance between the two nuclei increases, this description becomes less and less accurate, which demonstrates the importance of the UHF model, not only for systems with unpaired electrons, but those that would behave as such.

The differences between the RHF and UHF methods can be explained by

$$\chi(\boldsymbol{x}) = \begin{cases} \psi^{\alpha}(\boldsymbol{r})\alpha(\omega) \\ \quad \text{or} \\ \psi^{\beta}(\boldsymbol{r})\beta(\omega) \end{cases} \tag{1.83}$$

In the RHF model, it is required that $\psi^{\alpha}(\boldsymbol{r}) = \psi^{\beta}(\boldsymbol{r})$; however, this restriction is not imposed under UHF theory. This is possible by not restricting the contributions of each basis function, ϕ, to be the same for the ψ^{α} and ψ^{β} orbitals. This leads to a new form of the Roothaan-Hall equations known as the Pople-Nesbet equations[23] which are expressed as

$$\boldsymbol{F^{\alpha}C^{\alpha}} = \boldsymbol{S^{\alpha}C^{\alpha}E^{\alpha}} \tag{1.84}$$

$$\boldsymbol{F^{\beta}C^{\beta}} = \boldsymbol{S^{\beta}C^{\beta}E^{\beta}} \tag{1.85}$$

These equation are interdependent as the α-Fock matrix, $\boldsymbol{F}^{\alpha}$ depends on the set of ψ^{β} and vice-versa. In the event of a closed-shell system at or near the equilibrium geometry, the sets of coefficients, $\boldsymbol{C}^{\alpha}$ and $\boldsymbol{C}^{\beta}$, would be equal and this method would yield the same answer as the RHF method.

An alternative to the UHF method for open-shell systems is the restricted open-shell HF (ROHF) method[24] which employs doubly occupied spatial orbitals where possible before using singly occupied spin orbitals for any unpaired electrons. This approach is far less common than the UHF method due to its more complicated nature and lesser accuracy.[6]

1.9 Basis Sets

With the HF energy now expressed in terms of basis functions or atomic orbitals, a brief overview of basis sets and the basis functions they contain is necessary. From the Schrödinger equation solution for the hydrogen atom, it would appear that Slater orbitals would be a wise choice for basis functions for any atom. A Slater type orbital (STO)[25] is defined as

$$\phi^{\text{STO}}(\boldsymbol{r}) = (x - A_x)^l (y - A_y)^m (z - A_z)^n e^{-\alpha(\boldsymbol{r} - \boldsymbol{A})} \tag{1.86}$$

where $|\boldsymbol{r} - \boldsymbol{A}| = \sqrt{(x - A_x)^2 + (y - A_y)^2 + (z - A_z)^2}$ describes the position of an electron with respect to the nucleus, α is the Slater exponent which controls the breadth of its distribution, and the set of l, m, and n are the angular momenta in the x, y, and z directions, respectively. The sum of l, m, and n are equal to the orbital angular momentum which defines the shape of the orbital.

While Slater orbitals are very convenient to work with as results tend to converge with fewer numbers of Slater orbitals, they are very difficult to integrate.[11] This difficulty has led to Slater orbitals being used very rarely in computational chemistry and has led to the use of other types of basis functions. The most commonly used type for chemical systems are Gaussian functions. These Gaussian type orbitals (GTOs)[26] are defined very similarly to STOs and are expressed as

$$\phi^{\text{GTO}}(\boldsymbol{r}) = (x - A_x)^l (y - A_y)^m (z - A_z)^n e^{-\alpha(\boldsymbol{r} - \boldsymbol{A})^2} \tag{1.87}$$

The main issue with GTOs is that it takes many Gaussian functions to approximate a single Slater function, which leads to the need to utilize far more Gaussian orbitals to construct molecular orbitals.[11] This increase in the number of atomic orbitals leads to increases in computational time. However, the added expense of more atomic orbitals is less than the cost of the integral calculations involving Slater orbitals.

As l, m, and n control the angular momentum, these are specified for different types of orbitals. For example

Table 1.2: Angular momenta

(l,m,n)	Atomic Orbital Designation
(0,0,0)	s
(1,0,0)	p_x
(0,1,0)	p_y
(0,0,1)	p_z
(2,0,0)	d_{x^2}

Thus, the only unknown quantity in the definition of the GTOs is the α. Fortunately, there are many basis sets in the literature that have predefined values of α for different orbital types of various different atoms. While some basis sets are designed for all atoms on the periodic table, others are more specific to the more commonly used elements or for a group of atoms for which the basis set was intended.

1.9.1 Minimal Basis Sets

The smallest basis sets, which utilize a single basis function per atomic orbital in the given atom, are known as minimal basis sets. Thus, a basis set for hydrogen and helium would contain a single basis function to approximate the 1s orbital. Any

element from Li to Ne would use a set of 5 basis functions, one for each of the 1s, 2s, $2p_x$, $2p_y$, and $2p_z$ atomic orbitals. While these types of basis sets are quite rare, contracted minimal basis sets are used on occasion. In a contracted minimal basis set, a contracted set of Gaussian orbitals is used to approximate a single atomic orbital. For the H_2 molecule, we have two atomic orbitals, so a minimal basis set would require two basis functions. This can be achieved using the STO-3G basis set by[11]

$$\phi_1^{\text{con}} = d_1\phi_1^{\text{GTO}} + d_2\phi_2^{\text{GTO}} + d_3\phi_3^{\text{GTO}} \tag{1.88}$$

where d_i are the contraction coefficients and are predefined. The molecular orbital would then be given by

$$\psi_i = c_{1,i}\phi_1^{\text{con}} + c_{2,i}\phi_2^{\text{con}} \tag{1.89}$$

Thus, while the molecular orbital technically only consists of two basis functions with weighting coefficients requiring optimization, each basis function is comprised of 3 Gaussian functions with predefined contraction coefficients. The STO-nG basis sets are incorporated into most quantum chemical software packages[20–22] where n denotes the number of contracted Gaussians that are used to describe each atomic orbital. However, due to their minimal nature, they leave significant room for improvement.

1.9.2 Split-Valence Basis Sets

As core orbitals are not often involved in any interesting chemistry, when computational cost is a concern it seems reasonable to want to describe valence orbitals with

more accuracy than those in the core. This can be achieved using a split valence basis set.[11] Examples of these would be the Pople basis sets, X-YZG, or X-YZWG. X-YZG bases are known as split-valence double-zeta basis sets and X-YZWG are known as split valence triple-zeta basis sets, etc. To describe how these bases work, an X-YZG basis set (e.g. 3-21G) would utilize a contracted set of X Gaussian primitives to describe each core orbital. It would then use a contracted set of Y GTOs and a contracted set of Z GTOs to describe each valence orbital. These Pople basis sets are among the most commonly used basis sets in quantum chemical software packages. There are ways to improve these basis sets further through the addition of polarization and diffuse functions that will be discussed in the next section.

In addition to the Pople basis sets, there are numerous other split valence basis sets. These include the cc-pVXZ set developed by Dunning and coworkers,[27–30] the SVP, TZP, and QZP bases from Alrichs,[31] and the pc-n sets developed by Jensen et al.,[32–36] to only name a few.

1.9.3 Polarization and Diffuse Functions

As the exact form of molecular orbitals is not known, it is common to add functions of higher angular momentum to increase the flexibility of the basis set.[6] For instance, for a molecule such as methane, CH_4, one could add a set of d-orbitals to the basis set despite neither carbon nor hydrogen possessing occupied d-orbitals. These orbitals of higher angular momentum are known as polarization functions and are denoted in a few different ways. For polarization functions, it is common to treat heavy and light atoms differently. Light atoms refer to those in period 1 of

the periodic table (i.e. H and He), while heavy atoms refer to any other element on the periodic table. We previously discussed split valence basis sets of the form X-YZWG. One of the most common triple-zeta basis sets is the 6-311G basis. Adding polarization functions to this basis set can be expressed in a few different forms.

	Heavy Atoms	Light Atoms
6-311G(d,p)	1 set of d	1 set of p
6-311G**	1 set of d	1 set of p
6-311G(d)	1 set of d	None
6-311G*	1 set of d	None
6-311G(3df,2pd)	3 sets of d + 1 set of f	2 sets of p and 1 set of d

While newer literature tends to favour notations using letters specifying the types of polarization functions added, the asterisk notation may be observed in some older literature.[6] Any time that a single asterisk, or a single letter (or set of letters) are used in this notation, this means that the specified functions are added to heavy atoms and no additions are made to the light atoms in the system.

When adding polarization functions to a specific atom, they are always added in sets. Thus, the addition of a set of p-orbitals requires the addition of a p_x, a p_y, and a p_z orbital. While this may not seem noteworthy as it is what one might expect, the d-orbitals are slightly different. In chemistry undergraduate courses, students typically deal with a set of 5 d-orbitals, (d_{z^2}, $d_{x^2-y^2}$, d_{xy}, d_{xz}, and d_{yz}); however, in computational chemistry, we most often deal with a total of 6 d-orbitals, obtained from every Cartesian combination (i.e. d_{x^2}, d_{y^2}, d_{z^2}, d_{xy}, d_{xz}, and d_{yz}). Similarly, for f-orbitals, theorists typically deal with all 10 Cartesian f-orbitals as opposed to the set of 7 that are dealt with in experimental chemistry.[6]

Diffuse functions are simply basis functions that have small exponents, α.[37] This small exponent causes the the distribution of the function to be very wide, or diffuse. This feature makes diffuse functions useful for conjugated systems or more importantly, anions.

The notation for diffuse functions is dealt with much in the same way as polarization functions. One can add diffuse functions to both heavy and light atoms, or just to the heavy atoms. However, there is not as much flexibility in the addition of diffuse functions. In the case of the 6-311G basis set, one could add a set of diffuse functions to both the heavy and light atoms to yield the 6-311++G basis or simply to the heavy atoms in the case of the 6-311+G basis. A diffuse s-orbital and a set of diffuse p-orbitals are added to every heavy atom in the molecule, whereas a single diffuse s-orbital is added to each and every light atom in the system.

1.10 Correlated Methods

While the HF method typically determines absolute electronic energies to within 1% of the exact answer, that error is typically very important when it comes to the determination of any property of chemical reactions. Thus, trying to draw meaningful results from HF calculations is naive.

The main issue with the Hartree-Fock method lies in its neglect of what is referred to as Coulombic electron correlation. In essence, the HF method treats electron-electron interactions in an overly simplistic manner. An electron under the HF model does not 'see' the other electrons in the system, but instead, an average distribution of charge from the remaining electrons in the system. Based on

the Löwdin definition,[38] this Coulombic electron correlation energy, E^{corr}, or simply electron correlation energy is defined as:

$$E^{\text{corr}} = E_{\text{exact}} - E^{\text{HF}}_{\text{CBS}} \tag{1.90}$$

where $E^{\text{HF}}_{\text{CBS}}$ refers to the energy obtained from a HF calculation using a Complete Basis Set (or infinite basis set) and E_{exact} is the exact non-relativistic energy.

A significant portion of the field of quantum chemistry is dedicated to the development of different methods to accurately determine this correlation energy.[6] Much like mathematicians strive to determine the most decimal points of π, computational chemists strive for accurately determining the energy of the He atom.

There are two separate approaches to determine the correlation energy: implicit methods and explicit methods.[6] Explicit methods actually contain terms involving the distance between an electron pair, r_{12}, but are far less common for molecules. Implicit models are the only ones that will be discussed herein. Within implicit models, there are two separate approaches: Post Hartree-Fock Methods and Density Functional Theory (DFT). DFT was the only correlated method used throughout this thesis and will be discussed in far more detail than the Post HF methods.

1.10.1 Post HF Methods

As the name suggests, Post Hartree-Fock methods use the HF wavefunction and add corrections to it to try and determine the correlation energy. There are numerous Post HF methods, but the most commonly used are Møller-Plesset Perturbation

Theory (MPPT),[39] Configuration Interaction (CI),[40] and Coupled Cluster (CC) Theory.[41]

Each of these methods add components to the HF wavefunction by considering different configurations of electrons in the occupied and unoccupied (or virtual) orbitals. The main difference between these methods is how these other configurations are determined.

CCSD(T) which refers to Coupled Cluster Singles, Doubles, and iterative Triples, is regarded as the gold standard in computational chemistry and is used in Chapter 4 as a reference for the exact energy. Singles, doubles, and triples simply refers to how many electrons are excited to the virtual orbitals in each added configuration.

1.10.2 Density Functional Theory

Unlike the Post HF methods, density functional theory does not use the Hartree-Fock wavefunction; in fact, it doesn't use a wavefunction at all.[6] DFT, instead, uses only the electron density, $\rho(\boldsymbol{r})$, which is a 3-variable function ($\boldsymbol{r} = (r_x, r_y, r_z)$) as opposed to the $4N$ variables involved in the wavefunction, $\Psi(\boldsymbol{x}_1, \boldsymbol{x_2}, \ldots, \boldsymbol{x}_N)$.

Hohenberg and Kohn demonstrated that the exact ground state energy of a system, and thus, numerous other properties, can be determined from the electron density, $\rho(\boldsymbol{r})$.[42] This raises the question of why anyone would ever deal with the far more complicated $\Psi(\boldsymbol{x}_1, \boldsymbol{x_2}, \ldots, \boldsymbol{x}_N)$ if $\rho(\boldsymbol{r})$ can tell us the same information. The answer lies in the fact that while Hohenberg and Kohn demonstrated that the exact energy can be determined from $\rho(\boldsymbol{r})$, no one has yet to determine how. For this reason, there are numerous different DFT methods that have been developed over

the years in an attempt to accurately describe chemical systems. This section will describe the evolution of DFT from its Hohenberg and Kohn origins to where it is today.

Before we get into the math of DFT, let us first consider the name. As previously mentioned, DFT is a theory that deals with the density, $\rho(\boldsymbol{r})$. But what is a functional? Much like a function, a functional takes an input and outputs a number. However, while functions take a variable, x, as an input (i.e. $f(x)$), a functional, $F[f(x)]$, takes a function, $f(x)$, as an input in order to output a number. The square brackets denote a functional relationship. In the case of DFT, the input function is the electron density.

Much like in the case of HF theory, the DFT electronic energy (under the Born-Oppenheimer approximation[12]) can be broken down into three components as shown[6]

$$E_0[\rho_0] = \langle \hat{T}[\rho_0] \rangle + \langle \hat{V}_{en}[\rho_0] \rangle + \langle \hat{V}_{ee}[\rho_0] \rangle \tag{1.91}$$

where $\langle \hat{T} \rangle$, $\langle \hat{V}_{en} \rangle$, and $\langle \hat{V}_{ee} \rangle$ are the expectation values of the kinetic energy, electron-nuclear attraction, and electron repulsion operators, respectively. The subscript 0, denotes that these refer to properties of the ground state. Each of these components of the energy, much like the total energy, are functionals of the ground state electron density, $\rho_0(\boldsymbol{r})$.

The electron-nuclear attraction component is easily expressed in terms of the density. This operator is often expressed in terms of the external potential, $\nu(\boldsymbol{r_i})$, as

follows[6]

$$\hat{V}_{en} \equiv \sum_{i-1}^{N} \nu(\boldsymbol{r_i}) \tag{1.92}$$

where $\nu(\boldsymbol{r_i})$ is given by

$$\nu(\boldsymbol{r_i}) = -\sum_{A=1}^{M} \frac{Z_A}{r_{ia}} \tag{1.93}$$

The electron-nuclear attraction energy is then given by

$$\langle \hat{V}_{en}[\rho_0] \rangle = \langle \Psi_0 | \sum_{i=1}^{N} \nu(\boldsymbol{r_i}) | \Psi_0 \rangle = \int \rho_0(\boldsymbol{r}) \nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \tag{1.94}$$

where the indistinguishable nature of electrons allows for the simplification of expression to the final integral form. Using this new definition, we can rewrite the electronic DFT energy as

$$E_0[\rho_0] = \int \rho_0(\boldsymbol{r}) \nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \langle \hat{T}[\rho_0] \rangle + \langle \hat{V}_{ee}[\rho_0] \rangle \tag{1.95}$$

Here is where the problems with DFT begin. While we have a simple functional for the electron-nuclear attraction term, there are not simple forms for the kinetic energy and electron repulsion components. In 1965, Kohn and Sham devised a method, now known as the Kohn-Sham (KS) method or Kohn-Sham DFT, to determine the electron density, ρ_0, without first having a wavefunction, and how to determine the energy from that density.[43] In theory, KS DFT is exact; however, the unknown exchange-correlation functionals prevent this exactness.

KS DFT is developed around a reference system of non-interacting electrons that move throughout the external potential, $\nu_{ref}(\boldsymbol{r})$.[6] The external potential for the reference system is defined such that the electron density for the reference, $\rho_{\text{ref}}(\boldsymbol{r})$, is equal to the exact ground state density, $\rho_0(\boldsymbol{r})$. As the electrons do not interact with one another in the reference system, the Hamiltonian is given by

$$\hat{H}_{\text{ref}} = \sum_{i=1}^{N} \left[-\frac{1}{2}\nabla_i^2 + \nu_{\text{ref}}(\boldsymbol{r_i}) \right] = \sum_{i=1}^{N} \hat{h}_i^{\text{KS}} \tag{1.96}$$

where $\hat{h}_i^{\text{KS}}$ represents the one-electron Kohn-Sham Hamiltonian and is given by the terms in square brackets above. Much like the Hartree-Fock equations that were described previously, the spatial KS orbitals, ψ_i^{KS}, are the eigenfunctions of the very similar, Kohn-Sham equations[43]:

$$\hat{h}_i^{\text{KS}} \psi_i^{\text{KS}} = \varepsilon_i^{\text{KS}} \psi_i^{\text{KS}} \tag{1.97}$$

In this equation, as before, ε_i, represents the energy of the i^{th} molecular orbital. But unlike HF theory, KS orbital energies do not signify the amount of energy required to remove an electron from that orbital. These values are essentially meaningless. The physical significance of the orbitals themselves has been debated, but there is significant evidence demonstrating their utility, especially considering how similar they are to those obtained from the HF method.

Based on equation (1.95), $\langle \hat{T}[\rho_0] \rangle$ and $\langle \hat{V}_{ee}[\rho_0] \rangle$ still need to be determined in order to find the DFT energy. From the reference system, Kohn and Sham defined

these values as

$$\langle \hat{T}[\rho_0] \rangle = \langle \hat{T}_{\text{ref}}[\rho_0] \rangle + \Delta \langle \hat{T}[\rho_0] \rangle \tag{1.98}$$

$$\langle \hat{V_{ee}}[\rho_0] \rangle = \langle \hat{V}_{ee,\text{ref}}[\rho_0] \rangle + \Delta \langle \hat{V}_{ee}[\rho_0] \rangle \tag{1.99}$$

This equation states that the exact energy component is that of the reference system, $\langle \hat{A}_{\text{ref}}[\rho_0] \rangle$, plus some small correction for electron correlation, $\Delta \langle \hat{A}[\rho_0] \rangle$. Substituting these definitions into equation (1.95) gives

$$E_0[\rho_0] = \int \rho_0(\boldsymbol{r}) \nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \langle \hat{T}_{\text{ref}}[\rho_0] \rangle + \langle \hat{V}_{ee,\text{ref}}[\rho_0] \rangle + \Delta \langle \hat{T}[\rho_0] \rangle + \Delta \langle \hat{V}_{ee}[\rho_0] \rangle \tag{1.100}$$

These two corrections for electron correlation are typically combined into one term, known as the exchange-correlation energy, $E_{xc}[\rho_0]$. This simplifies (1.100) to

$$E_0[\rho_0] = \int \rho_0(\boldsymbol{r}) \nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} + \langle \hat{T}_{\text{ref}}[\rho_0] \rangle + \langle \hat{V}_{ee,\text{ref}}[\rho_0] \rangle + E_{xc}[\rho_0] \tag{1.101}$$

As the name suggests, this term contains corrections for not only correlation, but also exchange energy (which stems from the antisymmetric wavefunction) and any kinetic energy not accounted for in a system of non-interacting electrons.

In equation (1.101), there remain 3 components that are unknown. To define the first, $\langle \hat{T}_{\text{ref}}[\rho_0] \rangle$, the kinetic energy for the reference system, we must first define the electron density in terms of the KS orbitals, ψ_i^{KS}:

$$\rho_0(\boldsymbol{r}) = N \langle \Psi_{\text{ref}} | \delta(\boldsymbol{r} - \boldsymbol{r}') | \Psi_{\text{ref}} \rangle = \sum_{i=1}^{N} |\psi_i^{\text{KS}}|^2 \tag{1.102}$$

where $\delta(x)$ is the Dirac delta function and Ψ_{ref}, is the “wavefunction" for the reference. Wavefunction is in quotations as DFT does not deal with a wavefunction.

From this definition of the ground state electron density, the kinetic energy can now be determined. Realizing that much like HF, the KS “wavefunction" is a single Slater determinant, $\langle \hat{T}_{\text{ref}}[\rho_0] \rangle$ can be described by

$$\langle \hat{T}_{\text{ref}}[\rho_0] \rangle = -\frac{1}{2} \langle \Psi_{\text{ref}} | \sum_{i=1}^{N} \nabla_i^2 | \Psi_{\text{ref}} \rangle = -\frac{1}{2} \sum_{i=1}^{N} \langle \psi_i^{\text{KS}}(1) | \nabla_1^2 | \psi_i^{\text{KS}}(1) \rangle \tag{1.103}$$

The second term, $\langle \hat{V}_{ee,\text{ref}}[\rho_0] \rangle$, can be defined with respect to the density itself based on the classical equation for repulsion in an averaged electric field. This is given by

$$\langle \hat{V}_{ee,\text{ref}}[\rho_0] \rangle = \frac{1}{2} \int \frac{\rho_0(\boldsymbol{r_1}) \rho_0(\boldsymbol{r_2})}{r_{12}} \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} \tag{1.104}$$

Using these new definitions for the kinetic and electron repulsion energies for the reference system, equation (1.101) can be rewritten as

$$\begin{aligned} E_0[\rho_0] = \int \rho_0(\boldsymbol{r}) \nu(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} - \frac{1}{2} \sum_{i=1}^{N} \langle \psi_i^{\text{KS}}(1) | \nabla_1^2 | \psi_i^{\text{KS}}(1) \rangle \\ + \frac{1}{2} \int \frac{\rho_0(\boldsymbol{r_1}) \rho_0(\boldsymbol{r_2})}{r_{12}} \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} + E_{xc}[\rho_0] \end{aligned} \tag{1.105}$$

or using the definition of the external potential, as

$$\begin{aligned} E_0[\rho_0] = -\sum_{A=1}^{M} Z_A \int \frac{\rho_0(\boldsymbol{r})}{r_{iA}} \mathrm{d}\boldsymbol{r} - \frac{1}{2} \sum_{i=1}^{N} \langle \psi_i^{\text{KS}}(1) | \nabla_1^2 | \psi_i^{\text{KS}}(1) \rangle \\ + \frac{1}{2} \int \frac{\rho_0(\boldsymbol{r_1}) \rho_0(\boldsymbol{r_2})}{r_{12}} \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} + E_{xc}[\rho_0] \end{aligned} \tag{1.106}$$

The only remaining component to the energy is $E_{xc}[\rho_0]$. The exact form of this functional is not known and this is the reason why there are countless DFT methods that have been developed over the years. There have been numerous approaches proposed to accurately determine $E_{xc}[\rho_0]$, with the B3LYP functional[44] being amongst the most popular. The difficulty in DFT is that there is no systematic way to improve the accuracy as there is going from HF theory to a post-HF method. While one functional may work very well for a certain class of compounds, it may fail spectacularly for a different class. For this reason, DFT benchmark studies are very common to determine the best functional for a specific class of molecules for which accurate energies are known.[6] In this way, the top performing methods can then be applied to other molecules of that class for which energies or other properties are not known. For a good review of the existing functionals, the reader can consult the cited reviews.[45–47] The remainder of this chapter will focus on the different types of functionals that have been developed.

One of the first models used for the exchange-correlation functional was the Local Density Approximation (LDA) method. The LDA model is based on the theory of the uniform electron gas in which the charge density is equivalent at all points throughout the gas. Hohenberg and Kohn showed that for a system, such as where the electron density changes slowly with position, that $E_{xc}^{LDA}[\rho_0]$ is equal to[42]

$$E_{xc}^{LDA}[\rho_0] = \int \rho_0(\boldsymbol{r})\varepsilon_{xc}[\rho_0]\mathrm{d}\boldsymbol{r} \tag{1.107}$$

where ε_{xc} is the exchange-correlation energy density of each electron in a uniform

electron gas with density, ρ_0. This energy is typically broken up into separate exchange and correlation components as such

$$\varepsilon_{xc}[\rho_0] = \varepsilon_x[\rho_0] + \varepsilon_c[\rho_0] \tag{1.108}$$

A very simple form exists for the exchange component and it is given by

$$\varepsilon_x[\rho_0] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \rho_0(\boldsymbol{r})^{1/3} \tag{1.109}$$

For the LDA method, the correlation component is far more complicated. It was developed by Vosko, Wilk, and Nusair, and is denoted by ε_c^{VWN}.[48] Its functional form is as follows:[8]

$$\begin{aligned}\varepsilon_c^{VWN}[\rho_0] = \frac{A}{2}\Bigg[&\ln\left(\frac{x^2}{X(x)}\right) + \frac{2b}{Q}\tan^{-1}\left(\frac{Q}{2x+b}\right) \\ &- \frac{b\,x_0}{X(x_0)}\left(\ln\left(\frac{(x-x_0)^2}{X(x)}\right)\right) + \frac{2(b+2x_0)}{Q}\tan^{-1}\left(\frac{Q}{2x+b}\right)\Bigg]\end{aligned} \tag{1.110}$$

which consists of the following set of definitions:

$$\begin{aligned}&x = \left(\frac{3}{4\pi\rho_0(\boldsymbol{r})}\right)^{1/6}, \qquad X(x) = x^2 + bx + c, \qquad Q = (4c - b^2)^{1/2} \\ &A = 0.0621814, \qquad x_0 = -0.409826, \qquad b = 13.0720, \qquad c = 42.7198\end{aligned} \tag{1.111}$$

As seen here, the expression for ε_c is far from trivial. Both ε_x and ε_c are negative numbers; however, ε_x is typically much larger than ε_c. This fact is exploited by the

Xα method which omits the correlation component and scales the exchange energy by a factor, α, to account for this omission. While the correlation component is smaller than exchange, it is highly important as highlighted by the accuracy of the HF method. Further, it is not easily captured by scaling the exchange component. For this reason, the Xα method is rarely used today.

While the LDA model accurately describes metallic and carbon networks where the electron density does not change rapidly through space, it performs quite poorly for many other systems. Far more common are the functionals that use the generalized gradient approximation in full (GGA) or in part (meta GGA, hybrid GGA, hybrid meta GGA).

Unlike the LDA method, GGA methods are also concerned with changes in the electron density. For this reason, they are not only functionals of the density, but also of the gradient of the density, $\nabla\rho_0(\boldsymbol{r})$. Much like there exists restricted and unrestricted Hartree-Fock calculations, analogous treatments are available in DFT. Many GGA functionals are expressed in terms of an unrestricted treatment and thus involve the density for each of the α and β electrons ($\rho_0^\alpha(\boldsymbol{r})$ and $\rho_0^\beta(\boldsymbol{r})$). Thus, the general expression for the exchange-correlation energy for a GGA method is[8]

$$E_{xc}^{GGA}[\rho_0^\alpha), \rho_0^\beta] = \int f[\rho_0^\alpha(\boldsymbol{r}), \rho_0^\beta(\boldsymbol{r}), \nabla\rho_0^\alpha(\boldsymbol{r}), \nabla\rho_0^\beta(\boldsymbol{r})]\mathrm{d}\boldsymbol{r} \tag{1.112}$$

where f is a functional of the listed functions. The actual functional form varies depending on which GGA method is used. Some of the more accurate exchange functions have been developed by Becke[49,50] as well as Perdew and Wang[51,52]. For

instance, the B88 (also referred to as Bx88 and simply B) exchange functional has the form[49,50]

$$E_{xc}^{B88}[\rho_0^\alpha, \rho_0^\beta] = E_x^{LSDA} - b \sum_{\sigma=\alpha,\beta} \int \frac{(\rho_0^\sigma)^{4/3} \phi_\sigma^2}{1 + 6\, b\, \phi_\sigma \ln[\phi_\sigma + (\phi_\sigma^2 + 1)^{1/2}]} \mathrm{d}\boldsymbol{r} \tag{1.113}$$

where b is an empirical parameter equal to 0.0042 atomic units and

$$\phi_\sigma = \frac{|\nabla \rho_0^\sigma|}{(\rho_0^\sigma)^{4/3}} \tag{1.114}$$

The functional also contains the exchange energy from the LSDA method, E_x^{LSDA} which is the unrestricted version of the LDA method.[8] The energy expression is similar to that in (1.109). Purists tend to dislike the use of an empirical parameter in this functional; however, most commonly used functionals currently contain a number of empirical parameters.

As for the correlation functional, one of the most commonly used in GGA methods is the LYP functional which was developed by Lee, Yang, and Parr.[53,54] Others include the Pc86[55,56] and PWc91 (more commonly denoted simply as PW91), where P and W, once again, denote Perdew and Wang.[52] In theory, one can combine any exchange functional with any correlation functional; however, in practice there are common combinations that are used and these are incorporated into many quantum chemical software packages.[20–22] To name a given DFT method, the name of the exchange functional is typically combined with that of the correlation functional. For instance, a calculation that used the B88 or B exchange functional and the LYP ex-

change functional would be referred to as a B88LYP or, more commonly, a BLYP calculation.

In an attempt to improve upon GGA functionals, meta-GGA functionals were created. In addition to the gradient of the density, these functionals incorporate the second derivatives of ρ_0 and/or something referred to as the kinetic energy density, τ_σ. This property is defined as

$$\tau_\sigma = \frac{1}{2}\sum_i |\nabla\theta_{i\sigma}^{\text{KS}}|^2, \qquad \sigma = \alpha, \beta \tag{1.115}$$

where $\theta_{i\sigma}^{\text{KS}}$ is a KS spin orbital. The kinetic energy density can be incorporated into the exchange and/or correlation functional. A commonly utilized meta-GGA correlation functional is Becke's B95[57] which again, contains empirical parameters.

The final common type of functionals are the hybrid functionals.[6,8] There are hybrid GGA and hybrid meta-GGA functionals. These hybrids contain components from either hybrid GGA or meta-GGA functionals as well as from Hartree-Fock. Specifically, the exchange energy is borrowed from Hartree-Fock. This is commonly referred to as the exact exchange energy, E_x^{Exact}. For hybrid functionals, it is defined in terms of the KS orbitals

$$E_x^{\text{Exact}} = -\frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N}\langle\psi_i^{\text{KS}}(1)\psi_j^{\text{KS}}(2)|r_{12}^{-1}|\psi_j^{\text{KS}}(1)\psi_i^{\text{KS}}(2)\rangle \tag{1.116}$$

B3LYP, as previously noted, is by far the most commonly used functional in existence.[44] The method has been cited in the literature over 50,000 times since its

development in the early 1990s. The exchange correlation energy for the B3LYP method is given by[8]

$$E_{xc}^{\text{B3LYP}} = (1-a_0-a_x)E_x^{LSDA} + a_0\,E_x^{\text{Exact}} + a_x\,E_x^{B88} + (1-a_c)E_c^{VWN} + a_c\,E_c^{LYP} \quad (1.117)$$

The 3 in B3LYP corresponds to the 3 empirical parameters included in the expression, $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.81$. The remaining components of this equation have already been discussed earlier in this section.

As discussed, there is no functional that works better than all others. While hybrid methods tend to be more accurate than meta-GGA and GGA methods which tend to be better than the LDA method, this is not universally true. Benchmark studies for DFT are essential to determine which functional will work well on the class of molecules of interest. Throughout this thesis, for the purposes of accuracy, specific functionals were chosen for different applications based on previously conducted benchmark studies.

1.11 Electron Pair Descriptors

While Hohenberg and Kohn demonstrated that all properties could be determined from ρ_0, determining properties regarding electron pairs from a function that only deals with a single electron is far from trivial.[42]

Remember from earlier that the electron density can be obtained from the wavefunction by

$$\rho(\boldsymbol{r}) = N \int |\Psi(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \ldots, \boldsymbol{x_N})|^2 \mathrm{d}s_1 \mathrm{d}\boldsymbol{x_2} \mathrm{d}\boldsymbol{x_3} \ldots \mathrm{d}\boldsymbol{x_N} \tag{1.118}$$

Why not, instead, integrate over all but two electron position vectors, r_1 and r_2? This way, we obtain a function of two-electrons where extracting electron pair information is far simpler. This function, known as the pair density, $\rho(\boldsymbol{r_1}, \boldsymbol{r_2})$, is given by[6]

$$\rho(\boldsymbol{r_1}, \boldsymbol{r_2}) = \frac{N(N-1)}{2} \int |\Psi(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \ldots, \boldsymbol{x_N})|^2 \mathrm{d}s_1 \mathrm{d}s_2 \mathrm{d}\boldsymbol{x_3} \ldots \mathrm{d}\boldsymbol{x_N} \tag{1.119}$$

where $N(N-1)/2$ represents the number of pairs of electrons in the system. To rationalize this, consider the binomial coefficient or the permutation formula:

$${}_nP_k = \binom{n}{k} = \frac{n!}{(n-k)!k!} \qquad \therefore {}_NP_2 = \binom{N}{2} = \frac{N(N-1)}{2} \tag{1.120}$$

Unfortunately, the pair density, while far simpler than the wavefunction, is still too complex for visual representation. This suggests the need to simplify this function further. Consider the important properties of electron pairs. Things that come to mind are their separation, the position of their centre-of-mass, their relative velocities, etc. What do each of these properties have in common? They are all dependent on single scalar quantities that encompass electron pair information.

1.11.1 Intracules

Intracules represent a specific type of simplification of the pair density. They are probability densities that describe relative properties of electron pairs. For the purpose of this discussion, we will focus on 3 types of intracules: position, momentum, and posmom.

Position Intracule, $P(u)$

Let us first consider the easiest to conceptualize, the position intracule.[58] The position intracule, denoted as $P(u)$, describes the probability that any two electrons in a system will be separated by a distance $u = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$. The importance is of this function is obvious in that it describes interelectronic separations which are at the heart of repulsion between electrons. $P(u)$ can be obtained from the pair density as

$$P(u) = \int \rho(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta(u - |\boldsymbol{r_1} - \boldsymbol{r_2}|) \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} \mathrm{d}\Omega_u \tag{1.121}$$

Herein, Ω_u denotes the angular components of $\boldsymbol{u}$ and $\delta(x)$ is the 1-dimensional Dirac delta function. This function is equal to zero everywhere except $x = 0$. Furthermore, the following equality holds true for $\delta(x)$:

$$\int_{-\infty}^{\infty} \delta(x) \mathrm{d}x = 1 \tag{1.122}$$

This acts to define this new variable, u, as zero everywhere except where it is equal to separation between the two electron position vectors, $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$.

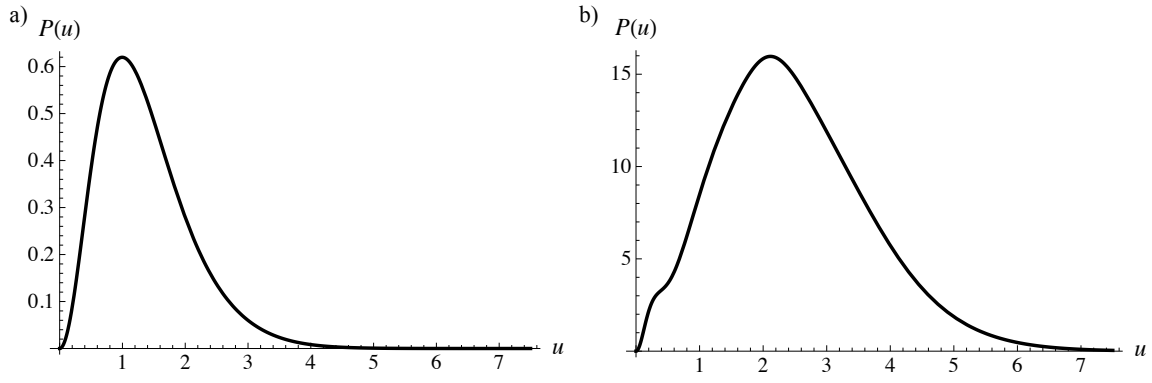


Figure 1.2: Position intracule for the ground state of a) the He atom, and b) the methane molecule.

Figure 1.2 depicts the position intracules for the ground states of the helium atom (left) and the methane molecule (right), both calculated at the HF level of theory. For helium, $P(u)$ is a simple unimodal distribution with a maximum at around 1.0 bohr or 0.53 Å. For a system as simple as helium with only 2 electrons, interpretation of $P(u)$ is rather simple as the function is describing the only electron pair in the system.

Considering the methane example now, one can see that the distribution is more complicated. While there is still a clear trend of an increased likelihood of electrons being further apart until a separation of approximately 2.2 bohr (1.16 Å), after which the probability decreases, it is a bimodal function. What leads to this much more complicated nature? Consider how many pairs of electrons exist in methane. As previously discussed when discussing the number of electron-electron interaction terms, any N electron system has $N(N-1)/2$ different pairs of electrons. For methane, a 10-electron system, this equates to 45 different pairs of electrons that are being described by $P(u)$ which explains why the function is more complex. The

issue with considering so many pairs of electrons is that some valuable information can be lost in all of the data. This is one of the main focuses of the present research and will be discussed further in the Project Goals section.

Of all probability distributions that describe electron pairs, the position intracule is the most studied. The reason for this is its inherent association to electron repulsion energies. In fact, the two-electron energy for any system is exactly determined by

$$E_{ee} = \int \frac{1}{u} P(u) \mathrm{d}u \tag{1.123}$$

Gill and co-workers have used the position intracule (amongst other types of intracules) to study the effects of electron correlation.[59–64] Intracule functional theory, a two-electron analogue of density functional theory, attempts to extract the correlation energy from an intracule by applying different correlation kernels. The effects of correlation will not be considered in great depth herein; however, previous research in the Pearson group has focussed on this.[65,66]

Momentum Intracule, $M(v)$

Thus far, the discussion in this chapter has focussed on position space. However, one can just as easily consider momentum space. Analogous to the pair density which describes the probability of finding electron 1 at $\boldsymbol{r_1}$ and electron 2 at $\boldsymbol{r_2}$, the two-electron momentum density, $\pi(\boldsymbol{p_1}, \boldsymbol{p_2})$ describes the probability of finding a pair of electrons with momenta $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$, simultaneously. This is obtained from

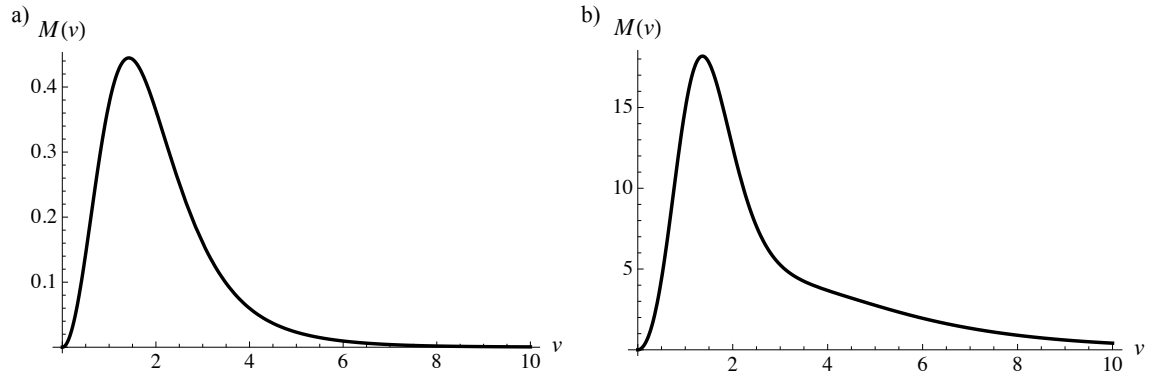


Figure 1.3: Momentum intracule for the ground state of a) the He atom, and b) the methane molecule.

the momentum space wavefunction, $\Phi(\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}, \ldots, \boldsymbol{p_N})$ as follows:

$$\pi(\boldsymbol{p_1}, \boldsymbol{p_2}) = \int |\Phi(\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}, \ldots, \boldsymbol{p_N})|^2 \mathrm{d}\boldsymbol{p_3} \ldots \mathrm{d}\boldsymbol{p_N} \tag{1.124}$$

The momentum space wavefunction, Φ, is the Fourier transform of the position space wavefunction, Ψ. Again, as intracules deal with relative properties, we are more interested in the scalar difference between $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ than the vectors themselves. The momentum intracule, $M(v)$, then, is obtained by[67]

$$M(v) = \int \pi(\boldsymbol{p_1}, \boldsymbol{p_2}) \delta(v - |\boldsymbol{p_1} - \boldsymbol{p_2}|) \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} \mathrm{d}\Omega_v \tag{1.125}$$

As before, the ground state momentum intracule of both the helium atom and methane are shown above (Figure 1.3). Unsurprisingly, the same type of behaviour is seen as before where $M(v)$ for the helium atom is a simple unimodal function while that for methane is again, more complicated. In both cases, the probability of the electrons having similar momenta is low as is the likelihood of v>5.

Posmom Intracule, $X(x)$

In the late 1990s, Rassolov suggested the importance of both u and v for proper understanding of electron correlation energies.[68] This problem was uniquely tackled by Bernard and Gill in their development of both the posmom density[69–71] and the posmom intracule.[1] Posmom refers to the dot product of the ***pos***ition and ***mom***entum vectors (for the posmom density) and their relative analogues (for the posmom intracule). For the purposes of this discussion, we will only focus on the intracule, $X(x)$. The origin of the posmom intracule (and in theory, all intracules) is the Wigner distribution.[72] Violating the Heisenberg Uncertainty Principle,[9] the Wigner distribution describes the simultaneous position and momentum of all electrons in the system. Although forbidden by quantum mechanics, there is nothing preventing the construction of this function, mathematically. It is determined from Ψ as follows:

$$W(\boldsymbol{r_1},\ldots,\boldsymbol{r_N},\boldsymbol{p_1},\ldots,\boldsymbol{p_N}) = \int \Psi^*(\boldsymbol{r_1}+\boldsymbol{q_1},\ldots,\boldsymbol{r_N}+\boldsymbol{q_N}) \times \Psi(\boldsymbol{r_1}-\boldsymbol{q_1},\ldots,\boldsymbol{r_N}-\boldsymbol{q_N}) e^{2i(\boldsymbol{p_1}\cdot\boldsymbol{q_1}+\cdots+\boldsymbol{p_N}\cdot\boldsymbol{q_N})} \mathrm{d}\boldsymbol{q_1}\ldots\mathrm{d}\boldsymbol{q_N} \quad (1.126)$$

There are a few notes regarding the Wigner distribution that should be discussed before proceeding further. First, it is not a true probability density as, although it is a real function, it does contain negative regions. For this reason, it is commonly referred to as a quasi-probability density. This does not, however, mean that it cannot be used to generate useful functions. Simplification of the function can be

achieved, as before, by averaging over all but a pair of electrons. This yields the second order Wigner distribution, $W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2})$, as

$$W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) = \int W(\boldsymbol{r_1}, \ldots, \boldsymbol{r_N}, \boldsymbol{p_1}, \ldots, \boldsymbol{p_N}) \mathrm{d}\boldsymbol{r_3} \ldots \mathrm{d}\boldsymbol{r_N} \mathrm{d}\boldsymbol{p_3} \ldots \boldsymbol{p_N} \quad (1.127)$$

An important property of this function to highlight is that it yields the exact pair density and two-electron momentum density:

$$\int W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) \mathrm{d}\boldsymbol{p_1} \mathrm{d}\boldsymbol{p_2} = \rho(\boldsymbol{r_1}, \boldsymbol{r_2}) \quad (1.128)$$

$$\int W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} = \pi(\boldsymbol{p_1}, \boldsymbol{p_2}) \quad (1.129)$$

Hence, both $P(u)$ and $M(v)$ can be obtained exactly from the second order Wigner distribution by substituting (1.128) into (1.121) and (1.129) into (1.125), respectively.

This demonstrates that the second order Wigner distribution contains all of the information that is present in both $\rho(\boldsymbol{r_1}, \boldsymbol{r_2})$ and $\pi(\boldsymbol{p_1}, \boldsymbol{p_2})$. Analysis of said function should thus, contain all information required for the posmom intracule, i.e. the dot product of $\boldsymbol{u}$ and $\boldsymbol{v}$. To transform this function into something with direction information regarding relative positions and momenta of electrons, $W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2})$ must first be transformed into the Omega intracule, $\Omega(u, v, \omega)$, as follows:

$$\Omega(u, v, \omega) = \int W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) \delta(u - r_{12}) \delta(v - p_{12}) \delta(\omega - \theta_{uv}) \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2} \mathrm{d}\boldsymbol{p_1} \mathrm{d}\boldsymbol{p_2} \quad (1.130)$$

The Omega intracule is the intracule from which all other intracules (with the exception of the posmom intracule) are derived. The dot intracule, $D(x)$, is a first order approximation to the posmom intracule. It is easily obtained from the Omega intracule by

$$D(x) = \int_0^\infty \int_x^\infty \frac{\Omega(u, z/u, \omega)}{u\, z \sin(\omega)} \mathrm{d}z\, \mathrm{d}u \tag{1.131}$$

where

$$z = u\, v \qquad \text{and} \qquad x = \boldsymbol{u} \cdot \boldsymbol{v} \tag{1.132}$$

As for the Posmom intracule, $X(x)$, it can be derived from the second order density matrix, $\rho_2(\boldsymbol{r_1}, \boldsymbol{r'_1}, \boldsymbol{r_2}, \boldsymbol{r'_2})$ as follows:

$$X(x) = \frac{1}{2\pi} \int \rho_2(\boldsymbol{r}, \boldsymbol{r} + \boldsymbol{u} \sinh k, \boldsymbol{r} + \boldsymbol{u}\, e^k, \boldsymbol{r} + \boldsymbol{u} \cosh k)\, e^{ikx} \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{u}\, \mathrm{d}k \tag{1.133}$$

However, when using the recurrence relation developed by Hollett and Gill (HG RR)[73], it has been show to be much easier to calculate the Fourier space analogue of $X(x)$. The Fourier space density, $\hat{X}(k)$, is given by

$$\hat{X}(k) = \int \rho_2(\boldsymbol{r}, \boldsymbol{r} + \boldsymbol{u} \sinh k, \boldsymbol{r} + \boldsymbol{u}\, e^k, \boldsymbol{r} + \boldsymbol{u} \cosh k) \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{u} \tag{1.134}$$

From a quick comparison of equations (1.133) and (1.134), it can be seen that once $\hat{X}(k)$ is known, performing an inverse Fourier transform will yield $X(x)$ as shown below:

$$X(x) = \frac{1}{2\pi} \int \hat{X}(k) e^{ikx} \mathrm{d}k \tag{1.135}$$

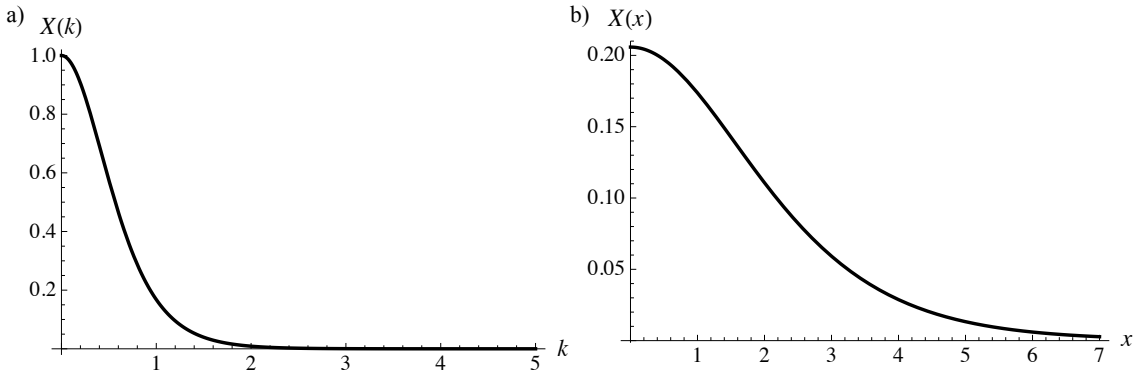


Figure 1.4: Posmom intracule for the ground state of the He atom in a) Cartesian space, and b) Fourier space.

In Figure 1.4, both $\hat{X}(k)$ and $X(x)$ are shown. These functions require more in-depth analysis to determine what the information is describing. Figure 1.5 below shows the different combinations of magnitudes for u and v and how the angle between them, ω, affects the resulting dot product, x. Nonetheless, the function contains more information than either the position intracule or momentum intracule on their own, and could provide the basis for strong predictions of experimental properties.

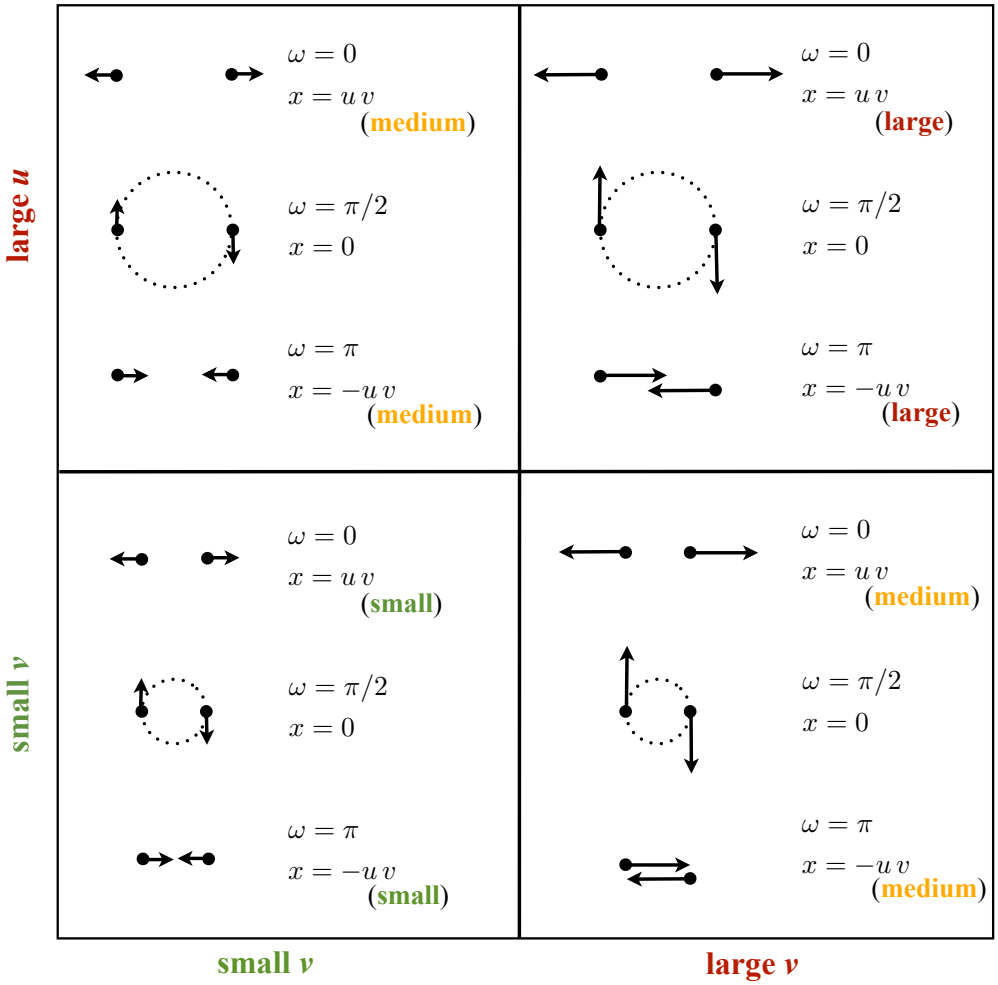


Figure 1.5: Possible combinations of u, v, and ω and the resulting values of $x = u \cdot v$. Adapted from the original version with permission from *Molecular Physics*.[1]

1.11.2 Extracules

Unlike the family of intracules, extracules contain *absolute* information instead of *relative* information. Whereas the position intracule describes the interelectronic separation of the electron pair, which is essential in determining electronic repulsions, the same cannot be said for the position extracule. It is, perhaps, for this reason that the family of extracules are far less studied than their intracular coun-

terparts. Nonetheless, they do contain valuable and complementary information. Throughout this research, there were 2 main types of extracules that were studied and that will be discussed in depth in this section: the position extracule, $E(R)$, and the momentum extracule $E(P)$.

Position Extracule, $E(R)$

The position extracule describes the probability that the centre-of-mass of an electron pair will be at a distance R from a predefined origin.[74,75] For atomic systems, this origin is typically defined as the nucleus, whereas for a molecular system, a number of different points could be considered depending on why the system was being studied. Much like $P(u)$, $E(R)$ can also be obtained from the pair density by

$$E(R) = \int \rho(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta(R - \tfrac{|\boldsymbol{r_1}+\boldsymbol{r_2}|}{2}) \mathrm{d}\boldsymbol{r_1}\, \mathrm{d}\boldsymbol{r_2}\, \mathrm{d}\Omega_R \tag{1.136}$$

While in the case of interelectronic separations, the angle of the vector isn't overly informative, this is far from true in the case of the extracular coordinate, $\boldsymbol{R}$. Thus, in addition to the scalar version, $E(R)$, the vectorized or 3-D form can also yield important information $E(\boldsymbol{R})$ and is perhaps even more useful. The 3-D extracule is obtained by

$$E(\boldsymbol{R}) = \int \rho(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta(\boldsymbol{R} - \tfrac{\boldsymbol{r_1}+\boldsymbol{r_2}}{2}) \mathrm{d}\boldsymbol{r_1}\, \mathrm{d}\boldsymbol{r_2} \tag{1.137}$$

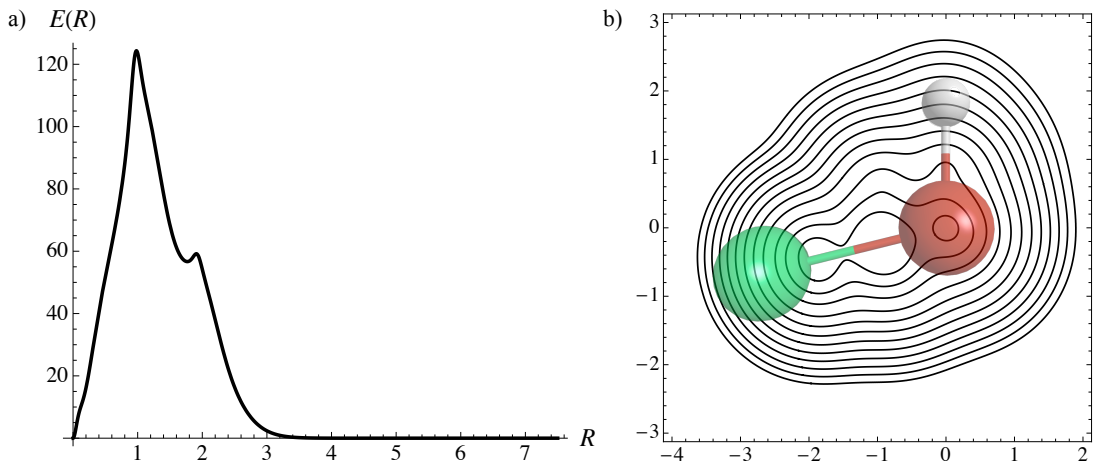


Figure 1.6: a) Scalar, $E(R)$, and b) vectorized, $E(\boldsymbol{R})$, position extracules for the ground state of HOF.

where $\delta(\boldsymbol{x})$ is now the 3-dimensional Dirac delta function. If $E(\boldsymbol{R})$ is known, $E(R)$ can be obtained through integration of the angular components of $\boldsymbol{R}$ as follows:

$$E(R) = \int_0^{2\pi} \int_0^{\pi} R^2 \sin\theta \, E(\boldsymbol{R}) \mathrm{d}\theta \, \mathrm{d}\phi \tag{1.138}$$

To highlight the differences between $E(R)$ and $E(\boldsymbol{R})$, both were calculated for the ground state of HOF and are shown in Figure 1.6. This molecule was chosen as it is one of the simplest asymmetric molecules with more than one bond. Atomic systems as well as those displaying high symmetry around the origin would not adequately demonstrate the differences between the vectorized and scalar forms of the position extracule. From $E(R)$, it can be noted that that the most probable values of R are 1 and 2 bohr. While one could theorize which pairs of electrons would result in these values for the electron centre-of-mass, the 3-D (or in this case, 2-D) position extracule can shed some light on the situation. The benefit provided by $E(\boldsymbol{R})$ only

grows as the molecule gets larger and the number of electrons continues to increase.

Momentum Extracule, $E(P)$

The momentum extracule, $E(P)$, is a probability density describing the average momentum, P, of an electron pair.[76] It can be obtained from the two-electron momentum density, $\pi(\boldsymbol{p_1}, \boldsymbol{p_2})$, in a similar fashion to the momentum intracule. Instead, we consider the Dirac delta function containing a descriptor for P as shown here:

$$E(P) = \int \pi(\boldsymbol{p_1}, \boldsymbol{p_2}) \delta(P - \tfrac{|\boldsymbol{p_1}+\boldsymbol{p_2}|}{2}) \mathrm{d}\boldsymbol{p_1}\, \mathrm{d}\boldsymbol{p_2} \tag{1.139}$$

An example of $E(P)$ is given below, again for the ground state of the helium atom and methane. The average momenta, P, has a smaller global maximum in the case of the larger methane molecule; however, due to the larger number of electron pairs, the probability of higher values of P are still significant. In comparison to the momentum extracule, the average momenta, P, displayed here are approximately half of the value of the relative momenta, v, shown in Figure 1.7.

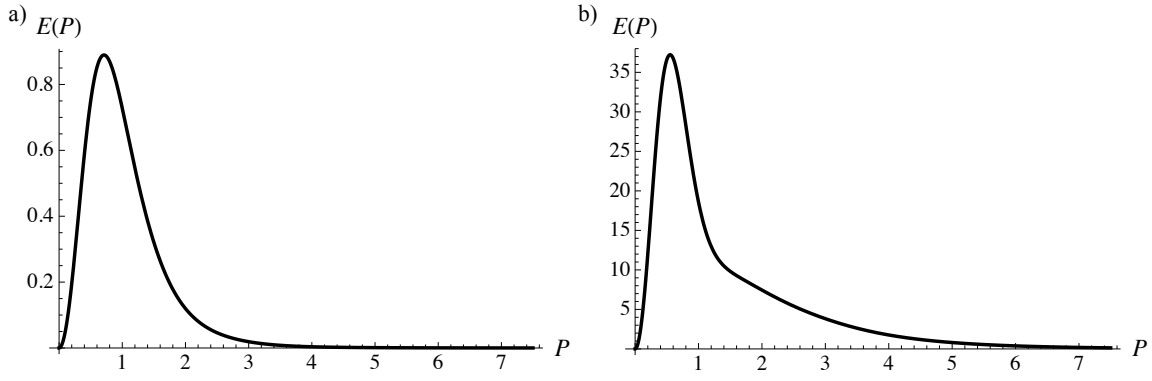


Figure 1.7: Momentum extracule for the ground state of a) the He atom, and b) the methane molecule.

1.12 Localized Molecular Orbitals

As previously noted, for systems with many pairs of electrons, it can be difficult to extract all relevant pieces of information from the intracule or extracule for the full molecule. While in theory, there are $N(N-1)/2$ pairs of electrons in a system, chemists typically only care about specific pairs, namely the $N/2$ pairs representing bonds, lone pairs, and to a lesser extent, core orbitals of the atoms within the molecule.

How then can we consider these $N/2$ pairs instead of every possible permutation of electrons in the system? The answer lies in the fact that the set of orbitals that are obtained by solving the Hartree-Fock or Kohn-Sham equations are not a unique solution. These orbitals are known as the canonical molecular orbitals, CMOs, and are delocalized over the full system. For any single determinant method, of which both HF and KS DFT are, a unitary transformation of the orbitals will not affect the molecular wavefunction. Thus, there exist other representations for molecular orbitals that will not affect the properties of the molecule as a whole. As the “wavefunctions" in these methods are based on a single determinant, let’s consider a simple example. If, for a two electron system, the wavefunction, Ψ is given by

$$\Psi = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc \tag{1.140}$$

Now, consider what happens if column 2 is added to column 1 (i.e. redefining orbital 1 as the sum of orbitals 1 and 2):

$$\Psi = \begin{vmatrix} a+c & c \\ b+d & d \end{vmatrix} = (a+c)d - (b+d)c = ad+cd - bc-cd = ad - bc \tag{1.141}$$

While the molecular orbitals do change, the wavefunction remains the same. Although this is an overly simplistic example that ignores corrections for normalization, it does demonstrate the ability to utilize different representations for the molecular orbitals without changing the properties of the system.

Lennard-Jones and Pople[77,78] were the first to suggest a method by which a set of localized orbitals that represent chemically intuitive orbitals (i.e. bonds and lone pairs) could be obtained. They noted that these orbitals could be obtained by minimizing the interorbital repulsions that are described by

$$4\sum_{i}\sum_{j>i}\int |\phi_i(\boldsymbol{r_1})|^2 \frac{1}{r_{12}} |\phi_j(\boldsymbol{r_2})|^2 \mathrm{d}\boldsymbol{r_1}\,\mathrm{d}\boldsymbol{r_2} \tag{1.142}$$

The 4 accounts for the four different interorbital interactions that occur between the electrons in a pair of orbitals.

This idea was considered by Edmiston and Ruedenberg (ER)[79] when they prepared the first method for obtaining localized molecular orbitals (LMOs). Other commonly used methods include that which was developed by Foster and Boys (Boys)[80] as well as the Pipek-Mezey (PM)[81] method. A comparison of the CMOs and LMOs of water is shown below in Figure 1.8.

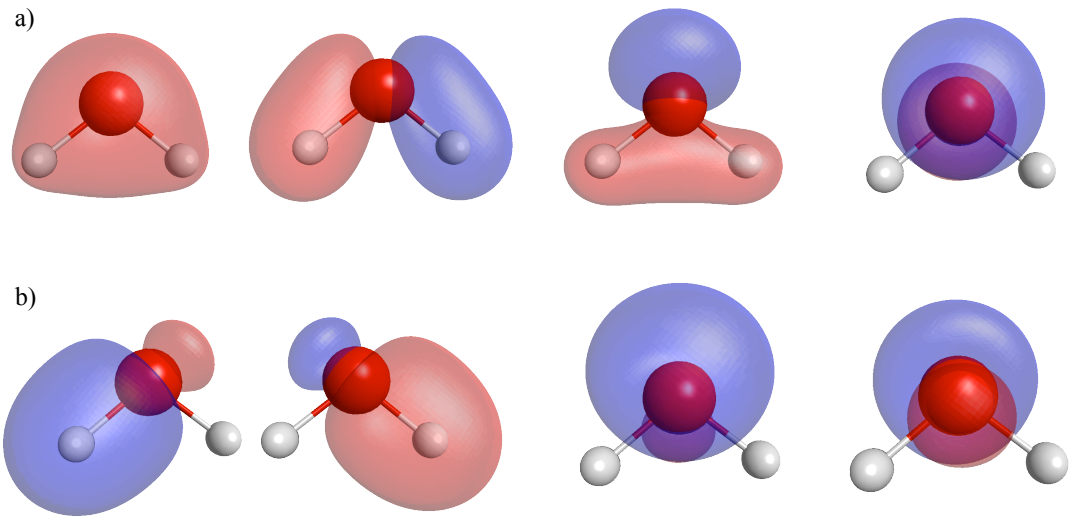


Figure 1.8: A depiction of the a) CMOs, and b) LMOs for the water molecule. Core orbitals are omitted.

Both of these methods tend to be much faster than the time-intensive ER method. The Boys LMOs tend to be very similar to those obtained from the ER method. The Boys method, unlike the ER method, is designed to maximize the squares of the distance between the centroids of charge of the occupied LMOs. The centroid of charge is defined as the point (x_c,y_c,z_c) where each of these coordinates is defined as

$$x_c = \langle\psi_i|x|\psi_i\rangle \qquad y_c = \langle\psi_i|y|\psi_i\rangle \qquad z_c = \langle\psi_i|z|\psi_i\rangle \tag{1.143}$$

Thus, the interpretation of the centroid of charge is the average position of each coordinate for the electron density of that orbital. If the distance between the centroids of charge is described by d_{ij}, the goal of the Boys localization method is to maximize

$$\sum_{i=1}^{N/2-1} \sum_{j>i}^{N/2} d_{ij^2} \tag{1.144}$$

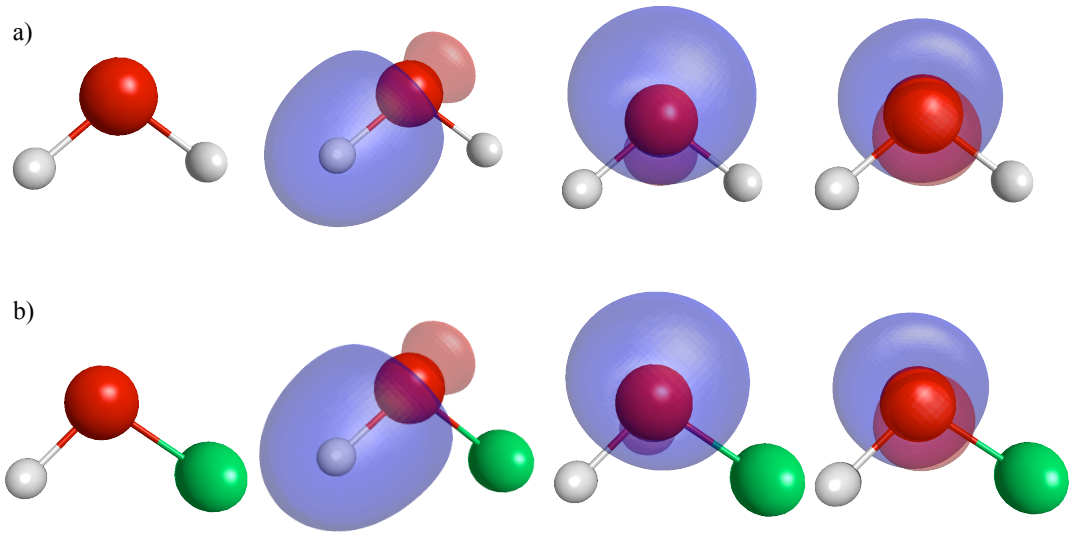


Figure 1.9: A depiction of the LMOs of a) H_2O, and b) HOF. Core orbitals are omitted.

Both of these methods treat double and triple bonds in a non-intuitive fashion. Instead of being comprised of σ and π bond(s), both the Boys and ER LMOs predict two (double bond) or three (triple bond) equivalent bonding LMOs that are equally spaced from one another.[6]

For a more intuitive treatment of covalent bonds of order greater than 1, the Pipek-Mezey method works quite well.[6] In the PM method, a double bond consists of a σ-bond LMO and a π-bond LMO. The main issue with the PM method is the representation of lone pairs. Unlike what one would expect for a lone pair orbital, those predicted by the PM method tend to resemble p-orbitals more than lone pairs.

One of the main benefits of LMOs is their transferability between different molecules. While HOH and HOF both contain an OH bond (two in the case of water), the set of molecular orbitals for these two molecules are significantly different. However, under the LMO representation, they are far more comparable (Figure 1.9).

Thus, by using these LMOs in place of their delocalized counterparts, it affords the opportunity to examine the effects on electron pair behaviour through the modification of the surrounding chemical environment.

When Lennard-Jones and Pople first theorized the development of LMOs, they discussed their potential to account for electron correlation. A set of LMOs should, in theory, represent a minimum of interorbital correlation making the intraorbital correlation the major component. LMOs are in fact utilized in a specific type of MPPT[39] calculations on larger molecules to speed up calculations.[6] Instead of using the CMOs to generate the HF reference, LMOs are used instead.

In their seminal paper on the extracule density, Thakkar and Moore[75] stated:

> *$E(\mathbf{R})$ deals with all the $N(N-1)/2$ electron pairs present in a N-electron molecule, but the chemist usually is interested only in $N/2$ 'chemical pairs'. [. . .] For RHF wavefunctions of systems with an even number of electrons, perhaps the 'chemical NSGs' can simply be taken to be det$|\phi_i\alpha\phi_i\beta|$ where the ϕ_i are localized molecular orbitals defined in one of the usual ways.*

While this statement was made in the context of extracules, it applies equally to any form of electron pair analysis. This approach of analyzing electronic structure with respect to localized molecular orbitals was first explored in our lab using the intracule density.[82,83] and will be expanded upon greatly throughout this thesis.

1.13 Project Goals

As previously alluded to, electron pair distributions can be rather difficult to interpret for larger systems. This is caused by the vast number of electron pairs ($N(N-1)/2$) in systems even as small as methane ($N = 10 \rightarrow 45$ pairs of electrons). This leads to potentially valuable information being lost in the immense amount of electron pair data.

Also, since chemists typically only care about bonds, and lone pairs, why consider every electron in the system? Why not consider a single pair of electrons? For instance, how does the electron pair in an OH σ-bond behave in the water molecule? How does that compare to its behaviour in HOF, in MeOH, or in EtOH. While this information wouldn't be readily accessible in the full electron pair distribution as it would be lost in the sea of information, it can be easily accessed by considering a single electron pair at a time. Due to the transferability of LMOs between molecules, it makes for the comparison of certain types of bonds or lone pairs between molecules trivial.

The goal of this thesis research was to provide a tool to answer these questions. The idea was to develop a software package that was capable of calculating each of the aforementioned electron pair distributions. This programming represents a significant portion of the research; however, one that won't be addressed in too much detail in this thesis as it is not of great interest to the field. Instead, this thesis will focus on the applications of this software package. While the potential applications of the software will only continue to grow as future users have new,

interesting questions that could be answered by this program, a few more general examples of the applications will be provided throughout this thesis. Much of the focus is on the underexplored position extracule density as previous work in the Pearson lab has focused primarily on the position intracule.

Chapter 2 will highlight the basics of the code and how it can be used to calculate these electron pair distributions (along with the User's Manual provided in Appendix A). Chapter 3 will present a broad overview of the potential applications of the software package while Chapters 4-6 will highlight specific examples, some of which were briefly addressed in Chapter 3. Finally, Chapter 7 will summarize the work and provide a vision for the future applications of this work.

Chapter 2

SEPDA software package

As mentioned in Section 1.12, a significant part of this research was the development of a software package capable of calculating intracules and extracules for a single electron pair. This software package is available at `https://j_pearson@bitbucket.org/aproud/sepda.git`. The purpose of this chapter is to describe the capabilities of the code and provide some details as to how it works. For more details about how to use the program, the reader can consult Appendix A.

2.1 Capabilities

While the intended purpose of the SEPDA package is to calculate electron pair properties for a localized molecular orbital, there are numerous other electron pair representations that can be considered. For instance, the natural bond orbital method devised by Weinhold and coworkers is an alternative localization procedure devised to provide the most accurate picture of a molecule based on the Lewis structure model. These NBOs can be calculated using the NBO software package, or through

the NBO add-on that is incorporated into more recent versions of the Gaussian software package.

Similarly, while CMOs are not localized, they do represent a description for a single electron pair. The SEPDA package is capable of dealing with all three types of orbitals: LMOs, NBOs, and CMOs. The focus of the research presented in this thesis, however, will be on the use of LMOs. CMOs present difficulty through their delocalized nature while NBOs, though popular, do not always possess occupations of 2.00 electrons.

2.2 Contents

SEPDA consists of three main components: the main code (sepda.csh), the main fortran programs, and the fortran subroutines. Additional components include files describing various predefined basis sets and numerous scripts in order to read in this information. A summary of the program architecture is given in Figure 2.1.

2.2.1 sepda.csh

The main program, sepda.csh, was written in C shell. A user prepares an input file for the program (as described in Appendix A) to tell the program what type of calculation is requested and where the necessary information for this calculation can be found. For instance, this input file would specify what type of orbital to analyze (LMO, NBO, or CMO), what type of electron pair density to calculate ($P(u)$, $M(v)$, $X(x)$, $E(R)$, $E(\boldsymbol{R})$, or $E(P)$), the basis set that was used, and where to find the MO coefficients.

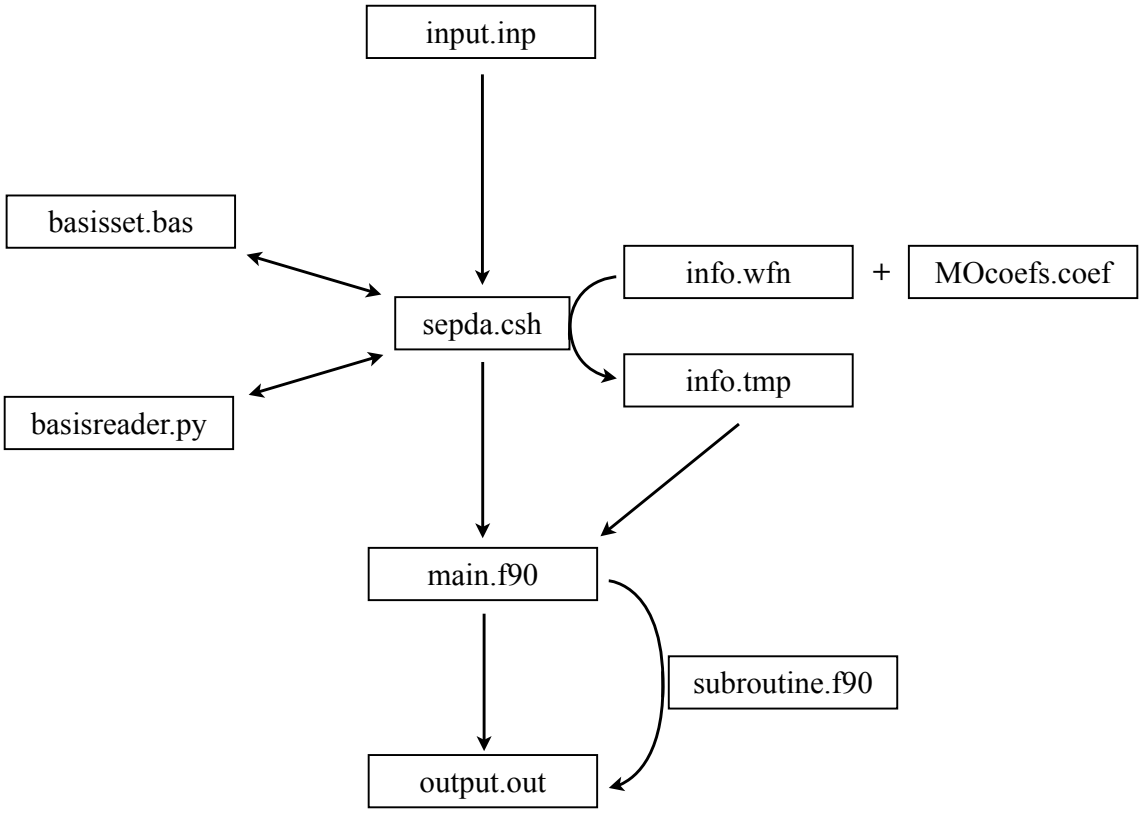


Figure 2.1: A general overview of the architecture of the SEPDA software package.

While it is most common to use pre-defined basis sets, a user has the flexibility to define their own basis set if desired. This program requires the specification of the basis set to determine whether it is pre-defined or user defined (BASIS=UD). In the case of a user-defined basis set, the program searches the input file to verify that the user did, in fact, provide a basis set definition (formatting is described in Appendix A). Using either the definition provided by the user or the pre-defined basis, this program can then verify that the number of basis functions expected is equal to the number of basis functions described by the MO coefficients. This takes into account any contraction of the basis sets and the contraction coefficients.

As the formatting of the MO coefficients for LMOs, NBOs, and CMOs all differ significantly, sepda.csh deals with each case separately and calls a script to read in the coefficients storing them in a temporary file in the appropriate format for reading by the main fortran program. All other necessary information is also copied to this file including any information regarding the basis functions (type, center, and exponent) as well as the coordinates of all atoms in the system.

Once all necessary information to calculate these electron pair distributions is known, the program reads in information regarding the calculation itself. This includes the type of calculation (($P(u)$, $M(v)$, $X(x)$, $E(R)$, $E(\boldsymbol{R})$, or $E(P)$)), the number of the molecular orbital of interest, and optional parameters regarding how dense the grid is for the calculation and how far from the origin this grid should extend.

Finally, using the information provided, sepda.csh calls the main fortran program and provides the necessary inputs in order to do the required calculation.

2.2.2 Main Fortran Programs

The main fortran programs (posInt.f90, momInt.f90, posmomInt.f90, posExt.f90, posvExt.f90, and momExt.f90) are based on earlier versions written by Zielinski[82] and then by Mackenzie for the position intracule, $P(u)$. For the purposes of this discussion, this code will be referred to as main.f90.

Based on the information provided by sepda.csh, the appropriate variant of main.f90 is called and supplied with information regarding the molecular orbitals. This information is read into a single mega array. Earlier versions of the code uti-

lized separate vectors/matrices for each piece of information. Comparisons were conducted on timing for these two separate approaches, and it was determined that small, but significant, enhancements were observed using the single mega array. This approach was then utilized for all variants of main.f90.

After this information is collected in the mega array, the MO coefficients are normalized except in the case of CMO coefficients which are output with normalization already included.

As the recurrence relation for these electron pair distributions calculates more than a single integral simultaneously, it is advantageous to have main.f90 break up the required integrals into classes. For instance, there are 21 separate classes for position space due to the 8-fold symmetry of the integrals:

[ssss]	[psss]	[ppss]	[psps]	[ppps]	[pppp]	[dsss]
[dpss]	[dsps]	[dsds]	[ddss]	[dpps]	[dspp]	[ddps]
[dpds]	[ddds]	[dppp]	[ddpp]	[dpdp]	[dddp]	[dddd]

In momentum space and combined posmom space, the 8-fold symmetry of position space is reduced to 4-fold symmetry. This results in the introduction of 6 new classes of integrals. Thus, in addition to the original 21 classes, the following 6 classes must be considered as well:

[pssp]	[dssp]	[dssd]	[dpsp]	[dpsd]	[dppd]

The number of s, p, and d type orbitals are determined based on the information provided in the .WFN file. From this, the number of different sets of p and d orbitals

can be determined. A single set of p type orbitals contains 1 p_x, 1 p_y, and 1 p_z while a set of d type orbitals contains 1 d_{x^2}, 1 d_{y^2}, 1 d_{z^2}, 1 d_{xy}, 1 d_{xz}, and 1 d_{yz}. The sets of d orbitals include all Cartesian possibilities. Thus, the number of sets of each can be determined as:

$$\text{pSets} = \frac{\text{pCount}}{3} \qquad \text{dSets} = \frac{\text{dCount}}{6} \tag{2.1}$$

where pCount and dCount are the total number of p and d type orbitals, respectively.

After determining the number of s type orbitals as well as pSets and dSets, it can be determined not only how many integrals of each class need calculation, but all combinations of basis functions that correspond to each class. With this information in hand, the recurrence relation can be called to calculate all the required integrals. These integrals can then be multiplied by the appropriate coefficients as follows:

$$\sum_{\mu\nu\lambda\sigma} c_\mu c_\nu c_\lambda c_\sigma (\mu\nu|\lambda\sigma)_x \tag{2.2}$$

where x corresponds to the specific type of electron pair distribution that is being calculated by main.f90. The summation goes over all integrals and this obtains a single point on the grid over which the distribution is calculated. This is repeated for every single point along the grid to obtain a series of points that can be interpolated to obtain the actual electron pair distribution of interest.

2.2.3 Fortran Subroutines - Recurrence Relation

As described in the last section, main.f90 makes calls the the appropriate subroutine where the integrals are actually calculated. For all scalar distributions, these integrals are calculated using the recurrence relation (RR) developed by Hollett and Gill (although the RR was adapted for the extracule calculations). The code for these calculations was also graciously provided by Hollett and Gill. As for the 3-Dimensional position extracule, these integrals were obtained using the recurrence relation developed by Thakkar and Moore. The code for this recurrence relation was written in house.

As much of this code was written by other researchers, a brief summary will simply be provided here. In these recurrence relations, all other integrals are obtained from the basic [0000] or [ssss] integral. The recurrence relation is then applied to this integral to obtain all others in iterations. The subroutine is broken down into each class of integral required for the specific electron pair distribution, and calculates the specific class of integrals based on the call from main.f90. For a derivation of the Hollett and Gill recurrence relations, the reader should consult Appendix B.

Chapter 3

Revealing Electron-Electron Interactions within Lewis Pairs in Chemical Systems

This chapter has been reproduced with modifications with permission from Proud, A.J.; Sheppard, B.J.H.; Pearson, J.K. *J. Am. Chem. Soc.* **2018**, *140*, 219-228. All of the work presented in this chapter was performed by Proud; however, some early work on position intracules for hydrogen bonding was conducted by Sheppard.

3.1 Introduction

It has now been over 100 years since Gilbert Lewis published his famous account of the electron pair in "The Atom and the Molecule" within this very journal.[84] In that article and a subsequent book[85], he eloquently laid out the details of his general theory of chemical bonding, from which he proposed that atoms form chemical

bonds by sharing *pairs* of electrons. A century later, the so-called "Lewis pair" is now entrenched in the *lingua franca* of modern chemical practice and the value of its predictive power is perhaps matched only by its remarkable simplicity. Chemists routinely employ Lewis's pioneering ideas about electron pairs to interpret molecular structure[86,87], intermolecular interactions,[88,89] property trends, and even the intricate mechanisms of chemical change.[90] Given the ubiquitous nature of the "Lewis pair" throughout chemistry, it is rather remarkable that we have not sought a definitive representation of such individual Lewis pairs that rigorously adheres to the quantum mechanical model of matter. Electron dot diagrams, as they are known, leave much to be desired in terms of quantitative information. To gain a more accurate understanding of electronic structure however, theoreticians will generally tout one of two alternatives for the fundamental variable in interpretive computations of electronic structure. There is, of course, the many electron wave function (Ψ), accounting for the spin and spatial variables of ***all*** electrons in the chemical system or the comparatively simple (yet no less informative) electron density (ρ), which usually describes the spatial *one-electron* probability distribution.[42] Neither Ψ nor ρ however, explicitly makes use of Lewis's notion of a localized electron pair in electronic structure. While the electron density does completely describe radical-based processes, these mechanisms do not possess the same level complexity as those involving the electron pair.

Highly intuitive questions are thus left unanswered in modern applications of electronic structure prediction. If we are so enamoured with the notion of several localized pairs of electrons dictating how we interpret molecular structure and

chemical behaviour, how is a single "Lewis" electron pair distributed in space? How fast do these electrons move relative to one another? How do such distributions differ from one pair to the next, and how might we exploit that information to gain insight into chemical systems, their properties, and the nature of chemical change? These are the questions we seek to answer herein.

To do so, we present a Single Electron Pair Distribution Analysis (SEPDA) technique that is both generally applicable and richly informative. We begin by defining a Lewis pair from a quantum mechanical perspective in terms of localized molecular orbitals.[82] Subsequently, we determine a series of interelectronic distribution functions for these orbitals from first principles and apply these techniques to a wide range of chemical contexts[83,91,92] including covalent bonding, non-covalent interactions, reaction coordinate diagrams, and more exotic 3-center 2-electron bonds, some of which will be addressed in more detail later in this thesis. We show that these distributions can yield highly useful and insightful information for interpreting electronic structure and predicting experimental behaviour.

3.2 Computational Methods

The SEPDA code, along with technical documentation and test cases is available from our code repository (`j_pearson@bitbucket.org/aproud/sepda.git`). The SEPDA program is designed to calculate the distribution of single electron pairs described by a given molecular orbital within position, momentum, or the combined "posmom"[1] space. As noted in the introduction, the starting point for an explicit discussion of electron pairs in position space is the spin-reduced two-electron den-

sity, $\rho_2(\boldsymbol{r}_1, \boldsymbol{r}_2)$, which is obtained from the molecular wave function by

$$\rho_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = \int |\Psi^*(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_N)|^2 \, \mathrm{d}s_1 \mathrm{d}s_2 \mathrm{d}\boldsymbol{x}_3 \ldots \mathrm{d}\boldsymbol{x}_N \tag{3.1}$$

To analyze electron pairs in momentum space, the spin-reduced two-electron momentum density, $\pi_2(\boldsymbol{p}_1, \boldsymbol{p}_2)$, can be obtained in an analogous fashion from the momentum space wave function by

$$\pi_2(\boldsymbol{p}_1, \boldsymbol{p}_2) = \int |\Phi^*(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \ldots, \boldsymbol{p}_N)|^2 \, \mathrm{d}\boldsymbol{p}_3 \ldots \mathrm{d}\boldsymbol{p}_N \tag{3.2}$$

However, since we are interested in the distribution of a *specific* pair of electrons, we must first determine the wave function for such a pair, which is conveniently achieved using any one of the well-known localization algorithms.[77–81,93–97] Though one has many choices when adopting a localization scheme (LMOs, NBOs, or even the delocalized CMOs) we have consistently employed the Edmiston-Ruedenberg (ER) localization technique for the work in this chapter. As previously noted, these so-called “chemically intuitive” orbitals quite naturally represent localized features of chemical structure such as bonds and lone pairs, and rigorously adhere to the quantum mechanical model of matter as they are obtained from a unitary transformation of the canonical molecular orbitals (CMOs). These are consequently an ideal starting point for our analysis and constitute our definition of a quantum mechanical “Lewis Pair" throughout this work. Constructing a two-electron determinant wavefunction in both position and momentum space from a single localized molec-

ular orbital (ψ_k) and substituting them in equations (3.1) and (3.2), respectively, yield the spin-reduced two-electron density in position, $\rho_2^k(\boldsymbol{r}_1, \boldsymbol{r}_2)$, and momentum space, $\pi_2^k(\boldsymbol{p}_1, \boldsymbol{p}_2)$, specifically for an electron pair described by ψ_k.

From these two-electron densities, we can obtain the desired electron pair distribution by applying the appropriate operator, $\hat{O}$, as shown in the following equations

$$\langle \hat{O} \rangle = \int \rho_2^k(\boldsymbol{r}_1, \boldsymbol{r}_2) \hat{O}(\boldsymbol{r}_1, \boldsymbol{r}_2) \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \tag{3.3}$$

$$\langle \hat{O} \rangle = \int \pi_2^k(\boldsymbol{p}_1, \boldsymbol{p}_2) \hat{O}(\boldsymbol{p}_1, \boldsymbol{p}_2) \mathrm{d}\boldsymbol{p}_1 \mathrm{d}\boldsymbol{p}_2 \tag{3.4}$$

The explicit dependence of these distribution functions on electron *pair* densities distinguishes SEPDA from other well-known localized electron density analysis techniques such as the electron localization function (ELF)[98] and the quantum theory of atoms in molecules (QTAIM).[99]

In the case of the relative separation of an electron pair, for example, $\hat{O} = \delta(u - |\vec{r}_1 - \vec{r}_2|)$ produces the position *intracule* density, $P(u)$, which, again, measures the probability of finding an electron pair separated by a scalar distance, u. The position *extracule* density describes the location of the centre-of-mass of a particular pair of electrons (R) and therefore allows us to track the motion of Lewis pairs using $\hat{O} = \delta(R - \frac{|\vec{r}_1 + \vec{r}_2|}{2})$.

Conversely, as disucssed, we may also calculate these quantities in terms of momentum, where the momentum intracule would yield the probability of observing a particular pair of electrons with a given relative momenta (v). The form of the operator required to obtain each type of analogous electron pair distributions are

defined below in Table 3.1. SEPDA has the ability to calculate all of these pair distributions for individual molecular orbitals: including the typical canonical molecular orbitals (CMOs), but also localized molecular orbitals (LMOs)[77–81], and natural bond orbitals (NBOs).[93–97] Of course, one should expect quite different results if using delocalized CMOs instead of chemically intuitive LMOs owing to their very different spatial distribution. These would necessarily require a different interpretation but it is of course true that the cumulative distribution of all possible electron pairs would be identical regardless of the orbital type used to construct the full wave function. It should be noted that the program is also capable of calculating these electron pair densities for the full molecule/atom, but this feature was added simply due to the simplicity of the required coding and is not novel.

In the table, the operator for the posmom intracule is denoted by *** as this probability distribution is obtained neither from the two-electron density nor the two-electron momentum density. Instead, it is obtained from the second order density matrix, $\rho_2(\boldsymbol{r_1}, \boldsymbol{r_1'}, \boldsymbol{r_2}, \boldsymbol{r_2'})$, as described by equations (1.134)-(1.136).

SEPDA employs recurrence relations to calculate the necessary two-electron integrals of arbitrary angular momentum for each of these pair distributions. While

Table 3.1: The scope of pair distributions available in the SEPDA software package.

Pair Distribution	Notation	$\hat{O}$	References
Position intracule	$P(u)$	$\delta(u - \lvert\vec{r}_1 - \vec{r}_2\rvert)$	[58, 59, 82, 83, 100–120]
Momentum intracule	$M(v)$	$\delta(v - \lvert\vec{p}_1 - \vec{p}_2\rvert)$	[67, 117, 119–141]
Posmom intracule	$X(x)$	***	[1]
Position extracule	$E(R)$	$\delta(R - \frac{\lvert\vec{r}_1+\vec{r}_2\rvert}{2})$	[74, 75, 100, 117–119, 142–148]
3-D Position extracule	$E(\vec{R})$	$\delta(\vec{R} - \frac{\vec{r}_1-\vec{r}_2}{2})$	[75, 91]
Momentum extracule	$E(P)$	$\delta(P - \frac{\lvert\vec{p}_1+\vec{p}_2\rvert}{2})$	[76, 117, 119, 131, 134, 135, 139, 140, 148]

the scalar densities utilize the recurrence relations for intracules developed by Hollett and Gill[73] with modified versions developed for the extracule counterparts, the 3-D position extracule employs the relation developed by Thakkar and Moore.[75] All of the optimized geometries and molecular orbitals utilized herein have been generated with the GAMESS suite of quantum chemistry programs. The level of theory employed in the current work to produce reliable structures and orbitals changes according to the chemical context of the applications. For example, when studying orbitals of isolated molecules it is perfectly reasonable to use a standard molecular orbital approach (i.e. Hartree-Fock theory) with a moderate basis set. In many cases however, such an approach is well known to fall outside the bounds of acceptable levels of error due to the absence of Coulomb correlation, which happens to be a highly relevant component when predicting most chemical properties. As such, Kohn-Sham density functional theory is arguably a more suitable starting point for generating structures and localized orbitals. Kohn-Sham DFT is well suited to this purpose as we have previously shown.[66] Additionally, Stowasser and Hoffmann have demonstrated that the Kohn-Sham orbitals can be reliably used to rationalize chemical phenomena and seem to be the orbitals a qualitative, chemical analysis needs.[149] As such, our choice of theoretical model for generating our structures and orbitals is governed by the chemical context and any relevant benchmark data available (*vide infra*).

For these calculations, the value of the desired pair distribution is determined at specific intervals along a grid defined by Mura and Knowles[150]. The breadth or scale of the grid as well as the number of grid points are user-defined with appropriate

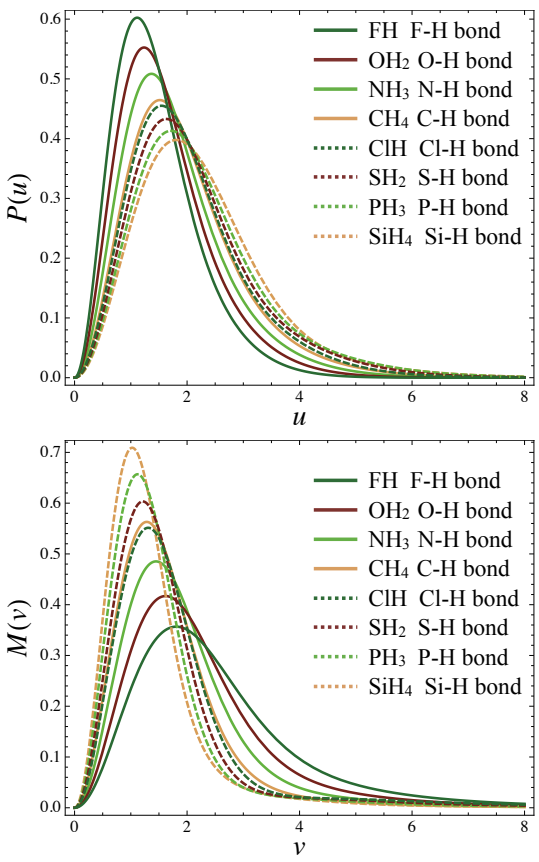
defaults established for each pair distribution type.

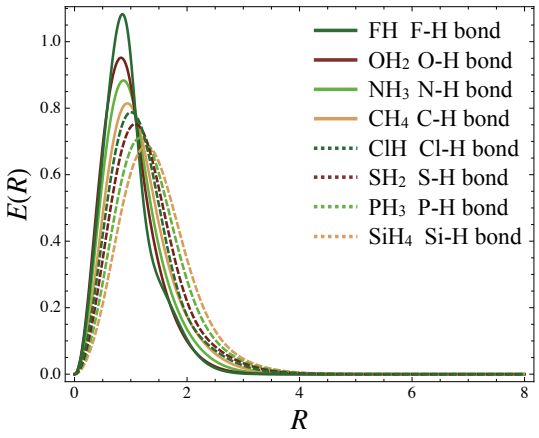
For 3-D position extracules, we have chosen specific planes within 3-D space for simplicity of presentation and analysis. In such cases, the molecule of interest is positioned appropriately in the yz-plane and instead of denoting this function as $E(0, R_y, R_z)$, we will simply use $E(\boldsymbol{R}_{yz})$. Atomic units are used throughout unless otherwise stated.

3.3 Results and Discussion

3.3.1 Covalent Bonding

We begin by considering one of the most fundamental of chemical features, the covalent bond. Figure 3.1 illustrates $P(u)$, $E(R)$, and $M(v)$ for X-H bonds in the set of first and second row hydrides as a simple model system to observe the effect on Lewis pair distribution of altering the heavy atom, X, within an X-H bond. As X is modified from the group 14 carbon to group 17 fluorine, we observe the bonding electron pairs *contracting* as they are drawn more strongly towards the heavy atom due to the associated increase in electronegativity. This same trend is seen with the second row hydrides.[82] Consequently, as these electrons are drawn closer together, they must move *faster* relative to one another, as is evidenced by the concomitant broadening of the momentum intracule, $M(v)$. Likewise, the centre-of-mass of the electron pair, R, migrates towards the heavy atom (which is positioned at the Cartesian origin), further confirming the nature of the contraction observed in $P(u)$. The SEPDA technique gives us both a qualitative and a quantitative picture of exactly





System	$\langle u^{-1} \rangle$	$\langle v^{-1} \rangle$	$\langle R^{-1} \rangle$	BDE (kJ/mol)
CH_4	0.6883	0.8257	0.8256	102.7
NH_3	0.7614	0.7215	0.7213	104.0
OH_2	0.8433	0.6325	0.6321	119.2
FH	0.9405	0.5532	0.5525	135.8
SiH_4	0.5893	1.0195	1.0193	80.0
PH_3	0.6101	0.9445	0.9442	88.5
SH_2	0.6381	0.8713	0.8709	91.0
ClH	0.6713	0.8024	0.8017	103.2
R^2	0.9476	0.9455	0.9457	- - -

Figure 3.1: Depiction of calculated $P(u)$, $E(R)$, and $M(v)$ for the first and second row hydrides at the HF/6-311G(d,p) level along with the first inverse moment, $\langle x^{-1} \rangle$ (where $x = u, R,$ or v) and experimental bond dissociation energies (BDE)[2,3].

how these bonding electron pairs are distributed and affords a unique opportunity to distinguish subtle but important chemical differences between these species.

As an interesting illustrative example, we have performed a linear least squares regression to show the strong predictive capacity that the distribution of bonding electron pairs have with respect to the experimentally determined strength of the bond itself. Shown in the figure are the first inverse moments, $\langle x^{-1} \rangle$ (where $x = u, R,$ or v), of each variable correlated with the experimentally measured bond dissociation energy of the X-H bond, where $\langle x^{-1} \rangle$ is determined by:

$$\langle x^{-1} \rangle = \int_0^\infty f(x)\, x^{-1} \mathrm{d}x \tag{3.5}$$

Moments of a function are a convenient and compact metric for broadly characterizing the distribution as a whole, and the first inverse moment is a natural choice in our case as $\langle u^{-1} \rangle$ corresponds *exactly* to the electron repulsion energy within the given orbital. From MO theory we would expect that an increase in the electronegativity of X would cause the HOMO to both be stabilized and more localized on the X atom. As such, one could expect to observe patterns in $P(u)$ (and consequently $M(v)$ and $E(R)$) that correlate to experimental BDE. Coefficients of determination, R^2, of approximately 0.95 were obtained for each of the three distributions. This indicates that by analyzing these pair densities for covalent bonds, we can gain significant insight into the strength of these specific interactions using any of the various tools available within the SEPDA package, and indeed this has been previously demonstrated in predicting acid dissociation constants as well.[83]

3.3.2 Non-Covalent Interactions

Extending our application of the technology to non-covalent interactions, we analyzed the water dimer using the breadth of tools available in the SEPDA package. In particular, we focused on the donor $\sigma_{\text{O-H}}$ bond and the acceptor n_{O} lone pair in the water dimer relevant for the H-bonding interaction, as shown in Figure 3.2. Unlike the hydrides where we simply calculated the pair distribution, here it is more instructive to present the *deformation* of the position intracule, $\Delta P(u)$, and the 3D

position extracule, $\Delta E(\vec{R}_{yz})$. These deformation densities are obtained as the difference between the relevant distribution for the hydrogen bonding complex and that of the isolated hydrogen bond donor (X-H) or acceptor (Y^-). The result indicates how the pair distributions *change* due to the formation of the hydrogen bond. In cases where the deformation densities are presented, we also indicate the *total content* of the deformation density, $\delta^k_{u,R}$, which is the integral of the magnitude of $\Delta P(u)$ or $\Delta E(\vec{R}_{yz})$ for a particular orbital k. This is a concise, scalar representation of the total difference between any two pair distributions.

Weinhold's resonance-covalency model[88,89] describes hydrogen bonding as resonance between

(3.6)

with the amount of the Y-H species dependent on the strength of the hydrogen bond. Based on this model, one expects a migration of the electrons from the n_O lone pair of Y^- to the hydrogen atom of X-H forming a weak 'covalent' bond between the oxygen and hydrogen (H-Y species). In doing so, the electron pair in the donor $\sigma_{\text{O-H}}$ bond of X-H would migrate towards the oxygen atom within that σ-bond. These two electron migration processes are indeed observed within $\Delta P(u)$ and $\Delta E(\vec{R}_{yz})$. First, for the electrons within the n_O lone pair, we see an increase in interelectronic separation from $\Delta P(u)$ and a shift in the electron pair centre-of-mass, $\vec{R}$, towards the hydrogen within the hydrogen bond donor species. Second, as the hydrogen bond forms we see a contraction of the σ-bond (X-H) electron pair from $\Delta P(u)$ and

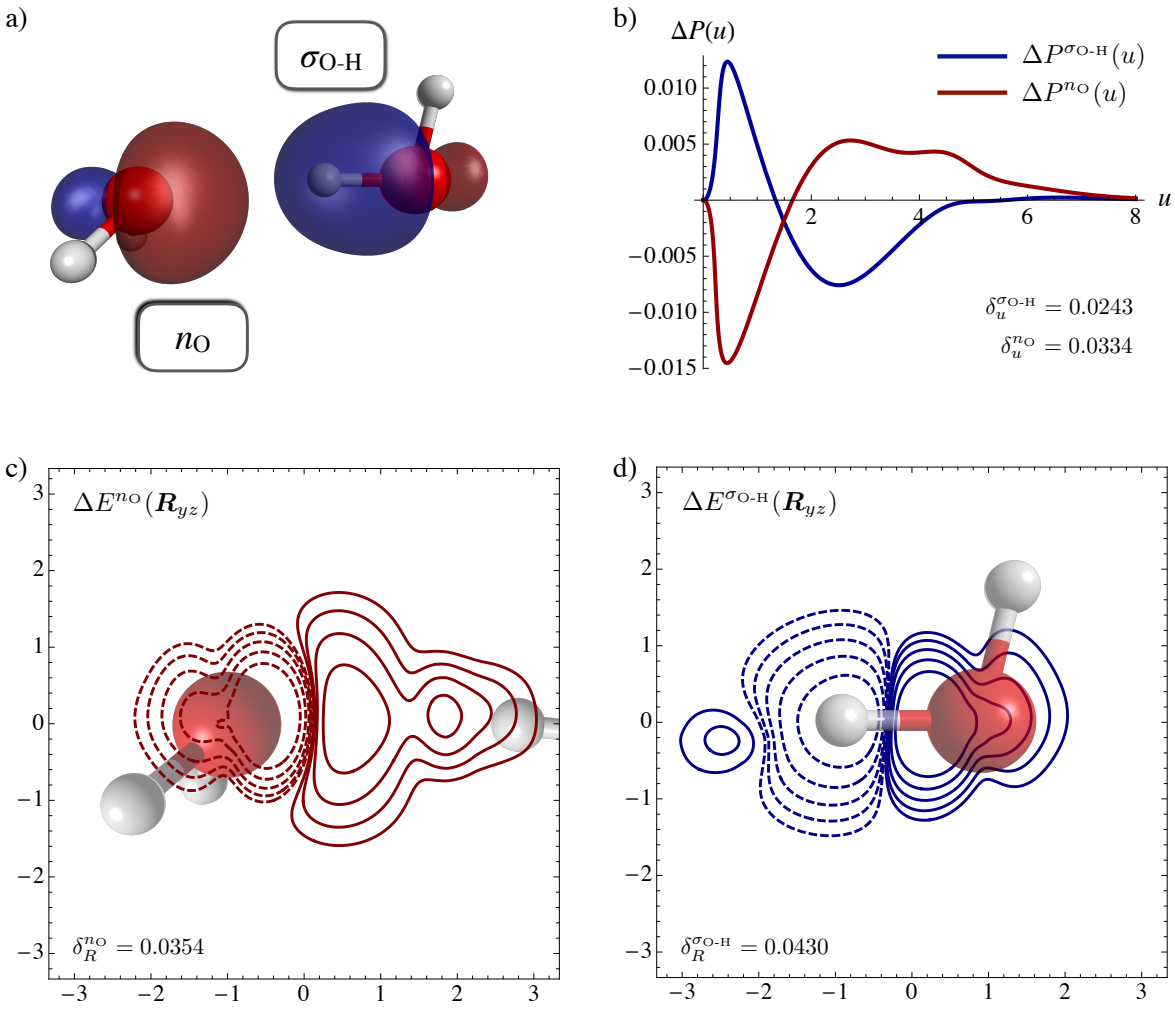


Figure 3.2: Depiction of a) the $\sigma_{\text{O-H}}$ and n_{O} ER LMOs of the water dimer calculated at the M06-2X/6-311G(d,p) level of theory, along with the b) $\Delta P(u)$ and c,d) $\Delta E(\vec{R}_{yz})$ for each. Solid and dashed lines in the $\Delta E(\vec{R}_{yz})$ plots denote positive and negative contours, respectively. Red contour lines correspond to $\Delta E(\vec{R}_{yz})$ for n_{O} and blue contour lines correspond to $\Delta E(\vec{R}_{yz})$ for $\sigma_{\text{O-H}}$. Contours are plotted for $\pm n \times 10^{-3}$ where $n = 4, 8, 16, 32, 64$.

a concomitant shift in the centre-of-mass of this electron pair towards the oxygen atom of the donor species from $\Delta E(\vec{R}_{yz})$. These two changes in absolute and relative electron pair positions agree very well with the resonance-covalency model of Weinhold and provide a far richer suite of quantitative information than a traditional Lewis interpretation.

We would like to highlight at this point that $\Delta E(\vec{R}_{yz})$ is an explicit *two-electron* distribution, as it maps the change in the distribution of the centre-of-mass of a localized electron pair. Though it resembles a traditional one-electron difference density, they are not directly related.

To further explore the effects of hydrogen bonding, we analyzed other hydrogen bonded complexes from Hobza's S66x8 data set.[151] In Figure 3.3, we show $\Delta P(u)$ and $\Delta E(\vec{R}_{yz})$ for the three hydrogen bonding systems: water-water, water-methanol, and water-methylamine. In each of these systems, the hydrogen bond donor is water, while the identity of the acceptor is systematically modified. However, when considering the three atoms directly involved in the interaction, we have: O-H-O (water-water), O-H-O (water-MeOH), and O-H-N (water-$MeNH_2$). In these first two hydrogen bonded complexes, the three atoms involved in the interaction are identical. Upon analysis of $\Delta P(u)$ and $\Delta E(\vec{R}_{yz})$ for these two systems, it is clear that they are very similar qualitatively. The main differences are seen in the extracule where the presence of the methyl group causes some distortions to the centre-of-mass of the electrons of the lone pair. In comparison to the water-methylamine system, significant differences are observed, especially in the intracule density. While the deformation density for the lone pair is very similar, there is an inversion of the relative heights of the maxima in the bimodal distribution at large u. In the case of the σ-bond, the deformation densities are very similar qualitatively; however, in the water-methylamine system, the magnitude of $\Delta P(u)$ increases by 40%. When one considers that of these three systems, water-methylamine has the strongest interaction energy,[151] this suggests the possibility of approximating inter-

action energies based on these deformation densities. Work in this area is currently underway in our research group on numerous hydrogen bonding systems within the S66x8 data set and the predictive properties have shown to be on par with or better than those obtained with bond critical points from Bader's Atoms in Molecules Theory.[152,153]

These techniques can be applied to other forms of non-covalent interactions as well, such as π-interactions, halogen bonding, etc. Figure 3.4 depicts $\Delta P(u)$ and $\Delta E(\vec{R}_{yz})$ for other types of non-covalent interactions: a) hydrogen bonding, b) halogen bonding, and c) π-interactions. While the hydrogen-bonded and halogen-bonded systems are very similar in nature, π-interactions are significantly different. In π-interactions, the electrons in the two π-orbitals donate electron density towards each other. The fundamentally different nature of this interaction is clearly visible in Figure 3.4. Both the position intracule and extracule are markedly different than those of the hydrogen and halogen bonding systems. They do however, exhibit the trends we would expect. $P(u)$ clearly shows that the relevant electron pair is separating when the π-interaction occurs and this is accompanied by a shift in the centre-of-mass towards the other ethene monomer within the dimeric system. The comparison of the results for the hydrogen and halogen bonding systems is interesting as these types of interactions are so similar. In both instances, a σ-bond with an uneven distribution of charge can lead to an interaction with a highly electronegative atom either inter- or intramolecularly. The key difference stems from the identity of the atom with a partial positive charge in the first molecule. In hydrogen bonding, this atom is hydrogen; however, in halogen bonding, the atom is

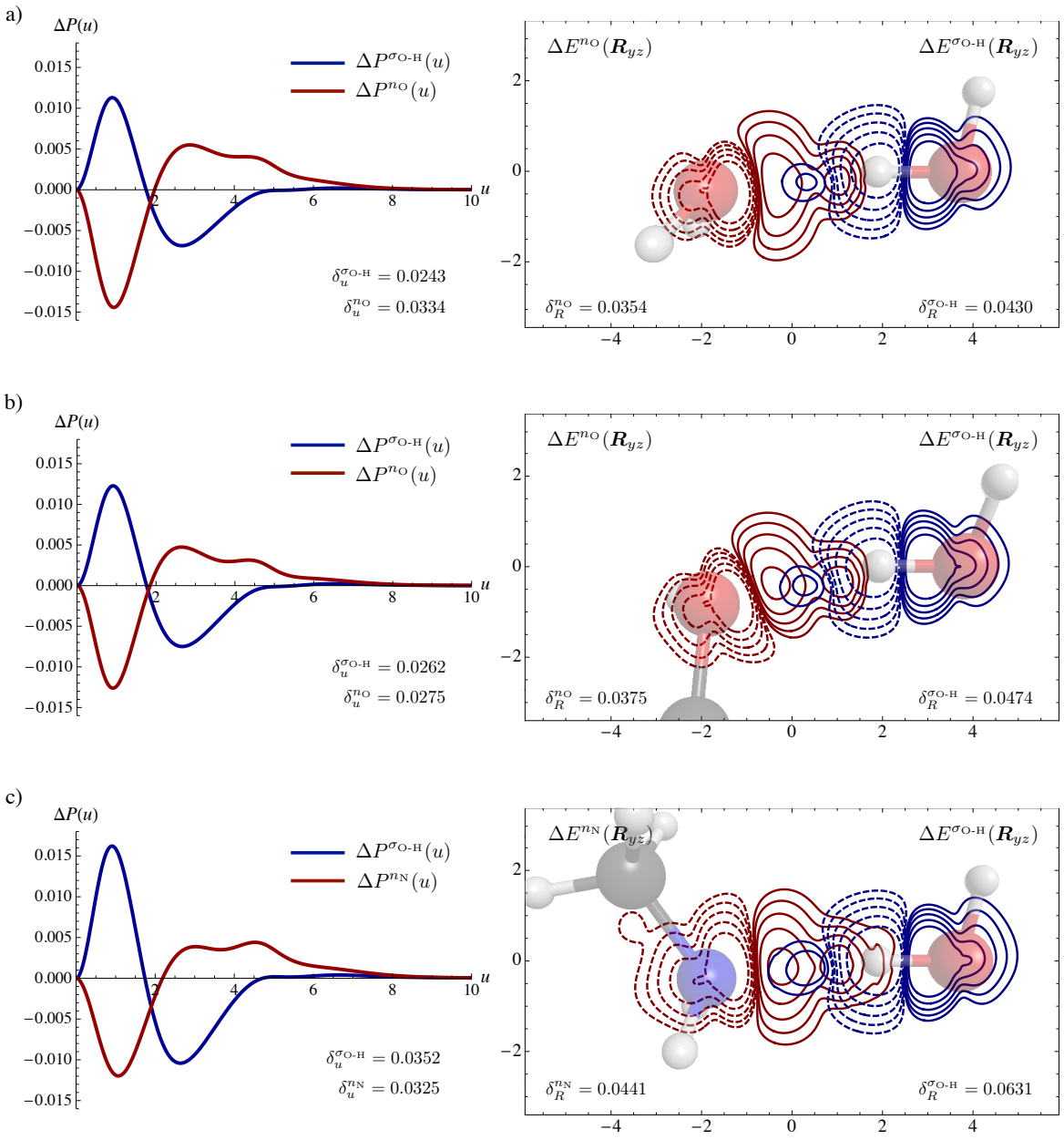


Figure 3.3: The deformation densities of the position intracules, $\Delta P(u)$, and 3-D position extracules, $\Delta E(\vec{R}_{yz})$, for the $\sigma_{\text{X-H}}$ and n_{Y} LMOs in the a) water dimer, b) water-methanol, and c) water-methylamine hydrogen bonding complexes. Solid and dashed lines in the $\Delta E(\vec{R}_{yz})$ plots denote positive and negative contours, respectively. Red contour lines correspond to $\Delta E(\vec{R}_{yz})$ for n_{O} and blue contour lines correspond to $\Delta E(\vec{R}_{yz})$ for $\sigma_{\text{O-H}}$. Contours are plotted for $\pm n \times 10^{-3}$ where $n = 4, 8, 16, 32, 64$. All data is calculated at the M06-2X/6-311G(d,p) level of theory.

one which is typically associated with partial negative charges: a halogen. However, in the presence of electronegative substituents with the halogen bond donor, the electron withdrawing substituents can reveal an electron-deficient region on the halogen atom known as the σ-hole.[154,155] It is this highly-localized electron-deficient spatial region which interacts with the electronegative atom of the acceptor to form the halogen-bonded interaction. Since the halogen of the donor also has electron rich regions, this results in differences in the intracule and extracules in comparison to hydrogen-bonding systems.

When analyzing $\Delta P(u)$ for these two systems, the overall trends are the same: the electrons of the σ-bonds contract while those of the lone pairs separate, again in accord with the resonance-covalency model of intermolecular interactions. However, more fine analysis of the topology yields some differentiating features. The deformation densities for the Y lone pairs do exhibit differences at large u though this may simply be due to the differences between n_{O} and n_{N}. The differences in $\Delta P(u)$ for the σ-bonds are quite significant as the maximum value for $\Delta P^{\sigma_{\mathrm{O\text{-}H}}}(u)$ is roughly half of that for $\Delta P^{\sigma_{\mathrm{F\text{-}Br}}}(u)$, suggesting that the differences are much less likely due to the identities of the atoms. Furthermore, in hydrogen bonding, $\Delta P(u)$ for the σ-bond is very simple, being unimodal both above and below the axis. In the halogen-bonding complex, it is primarily unimodal above the axis; however, at negative values of $\Delta P(u)$, the function is bimodal. This bimodal distribution is likely due to the more complicated nature of the atom with the partial positive charge. While the intracules can help to distinguish between these three different types of non-covalent interactions; only the π-interaction system demonstrates distinguish-

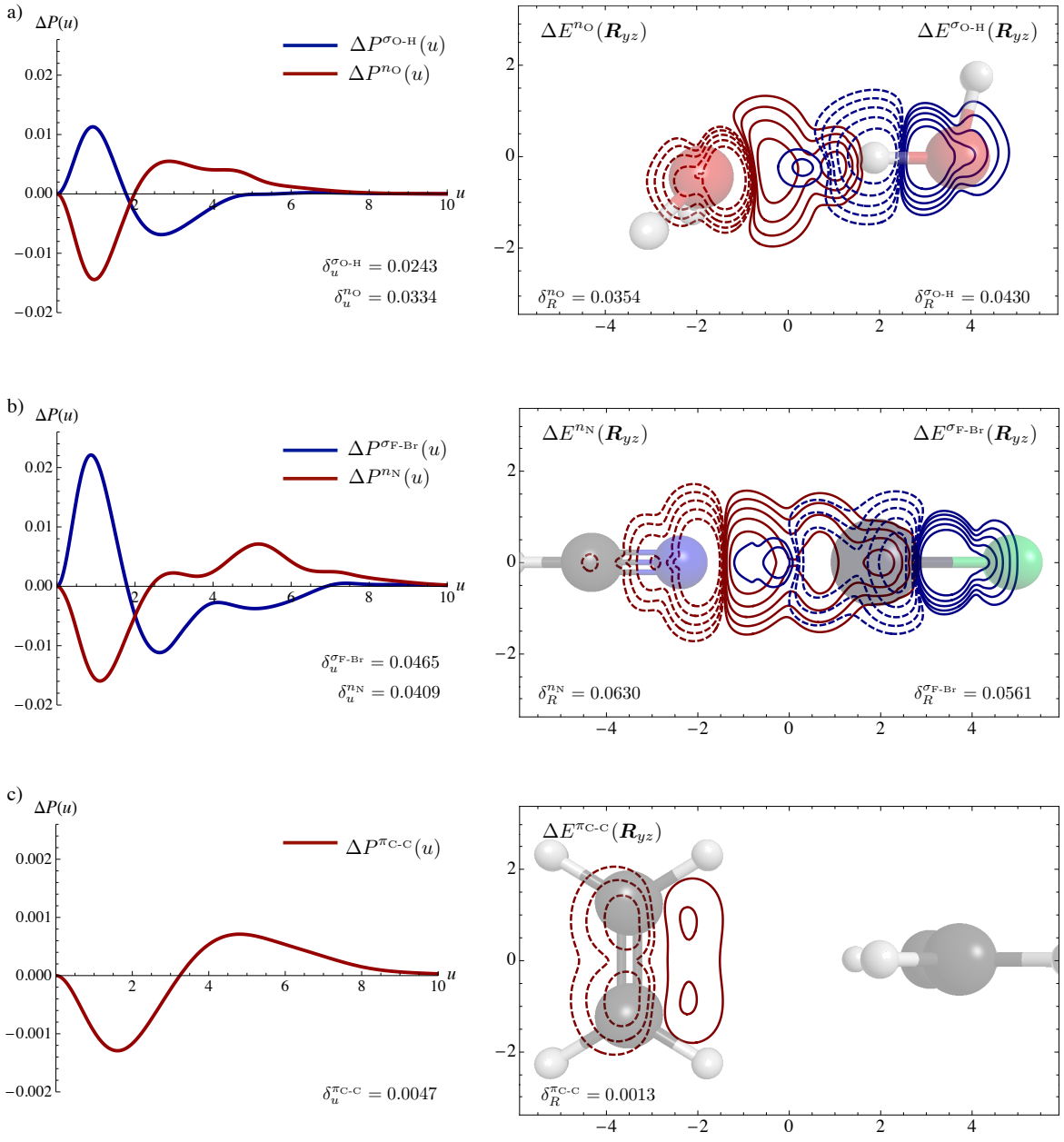


Figure 3.4: The deformation densities of the position intracules, $\Delta P(u)$, and 3-D position extracules, $\Delta E(\vec{R}_{yz})$, for the participating LMOs in the a) water dimer hydrogen bonding complex, b) FBr-HCN halogen bonding system, and c) the ethene dimer, π-interaction system. Solid and dashed lines in the $\Delta E(\vec{R}_{yz})$ plots denote positive and negative contours, respectively. Contours are plotted for $\pm n \times 10^{-3}$ where $n = 4, 8, 16, 32, 64$. Red contour lines illustrate $\Delta E(\vec{R}_{yz})$ for the non-bonding electron pair in a) and b) as well as the π bonding electrons in c). Blue contour lines indicate the σ bonding pair of electrons in parts a) and b). All data is calculated at the M06-2X/6-311G(d,p) level of theory.

able features for $\Delta E(\vec{R}_{yz})$. Regardless, the wealth of tools within the SEPDA package has the ability to quantitatively and qualitatively distinguish between different types of non-covalent interactions.

3.3.3 3-centre 2-electron Bonds

In addition to typical covalent bonds, we decided to study the distribution of Lewis pairs in more exotic species such as the 3-centre 2-electron bonds present in the B_2H_6 complex. Diborane represents a unique bonding environment, and as shown in Figure 3.5 our results for this system are rather interesting.

Based on the chemical environment, one would expect that the behaviour of the bridging hydrogens would be markedly different from that of the more typical terminal B-H bonds. However, the $P(u)$ for both the $\sigma_{\text{B-H-B}}$ and $\sigma_{\text{B-H}}$ bonds are nearly indistinguishable (Figure 3.5). There are only minor differences at large u that are noticeable by visual inspection. This is remarkable, considering the significant differences observed in the covalent bond intracules (above) for bonds that would only be considered marginally different. Because the B-H bond distance for the bridging hydrogens is 1.321 Å, while that of the terminal B-H bond is only 1.204 Å, and because the $\sigma_{\text{B-H-B}}$ spans three centres (and 1.750 Å between the two boron atoms), one would expect a much greater difference in the interelectronic separations of the electron pairs within the two orbitals. The extracule density however was indeed significantly different for the two bond types and provides evidence as to why the intracules are so similar. Despite the qualitatively different extracules, the spatial extent of the area covered by each is not drastically different. In other words, the

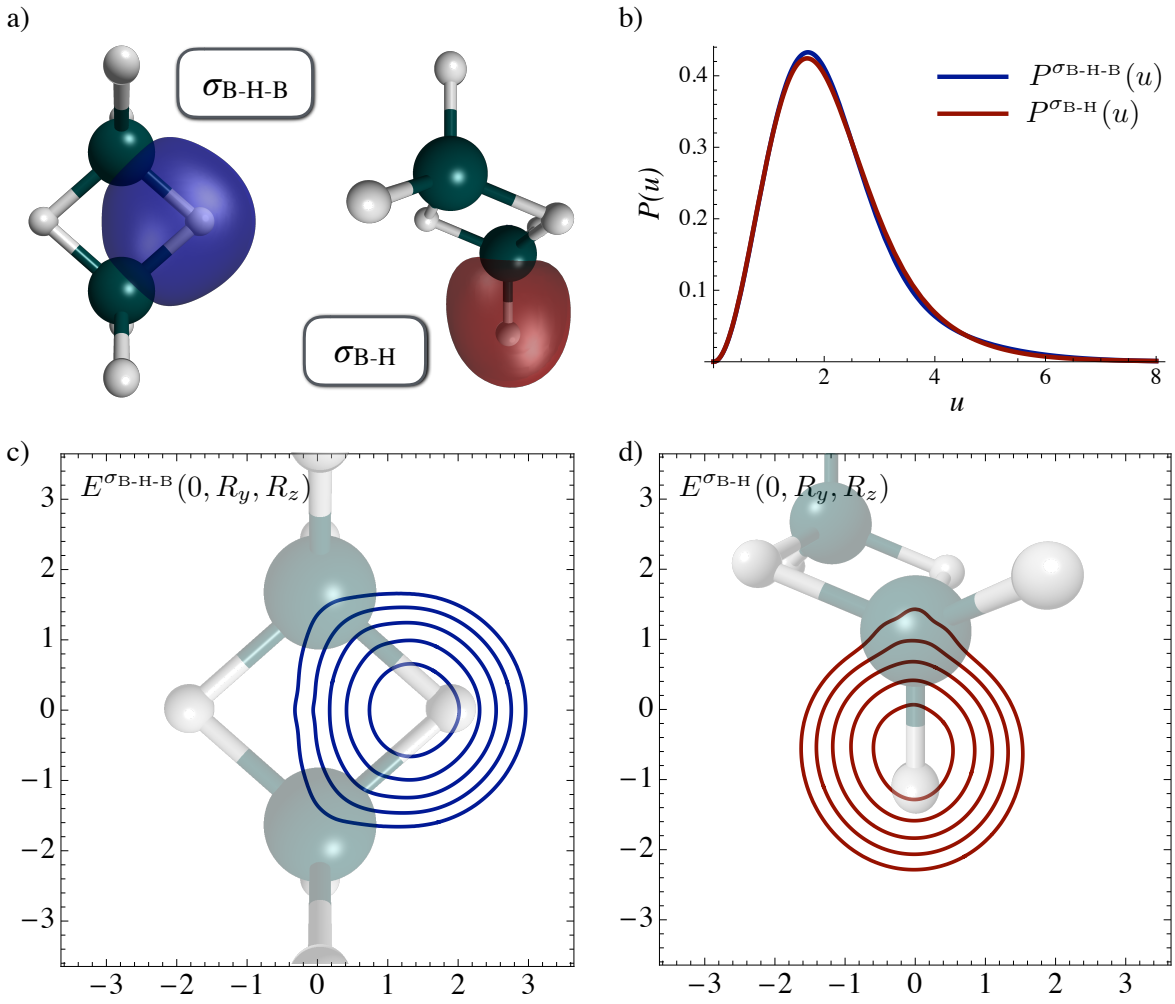


Figure 3.5: Depiction of a) the $\sigma_{\text{B-H-B}}$ and $\sigma_{\text{B-H}}$ LMOs of diborane with the associated b) $P(u)$ and c,d) $E(\vec{R}_{yz})$. Contours are plotted for $E(\vec{R}_{yz})$ values of 0.01, 0.02, 0.04, 0.08, and 0.016 atomic units. All data is calculated at the HF/6-311G(d,p) level of theory.

$\sigma_{\text{B-H}}$ and $\sigma_{\text{B-H-B}}$ localized electron pairs are distributed throughout a spatial region that is similar in size. This would allow for similar interelectronic separations, and means that the $\sigma_{\text{B-H-B}}$ pair is somewhat more localized to the inner bridging region than one might initially expect. This interesting result is supported by the relative weakness of the $\sigma_{\text{B-H-B}}$ bond relative to the $\sigma_{\text{B-H}}$ bond and provides evidence as to why the intracules are so similar.

3.3.4 Interpreting Reaction Mechanisms

Monitoring the progress of chemical reactions is often of significant interest when modelling chemistry and Lewis diagrams are a ubiquitous tool for chemists to intuit how reactants proceed to transition states and products via the reaction mechanism. Again though, SEPDA is an invaluable tool to gain additional qualitative and quantitative insight into how Lewis pairs migrate throughout a chemical reaction. An "arrow pushing" representation of a mechanism generally presumes that an electron pair is being spread over two atoms or condensed to one as the reaction progresses and the explicit two-electron treatment from SEPDA is a unique and intuitive tool to employ. For example, one could easily track the migration of a centre-of-mass of the pair with $E(R)$ while testing to see whether the inter electronic separation is increasing/decreasing with $P(u)$.

Utilizing an intrinsic reaction coordinate (IRC) calculation, whereby the electronic structure of a reaction complex is determined at a series of points from reactants to the transition state to products, one can employ the SEPDA package to monitor the change in the distribution of Lewis pairs throughout the course of the reaction. To demonstrate this application, consider the theoretical S_N2 reaction:

$$Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$$

$P(u)$ and $E(\vec{R}_{yz})$ were calculated at various points along the reaction coordinate between the reactant state and the transition state. Structures and energies were

calculated at the OLYP/6-311G(d,p) level of theory based on the benchmark studies performed by Bento *et al.*[4] Though it is certainly true that a single-determinant Kohn-Sham DFT model for non-equilibrium structures along a reaction coordinate will not yield the same degree of accuracy as other multi-determinantal model chemistries designed to capture static correlation effects, DFT has been readily applied to model reaction coordinate profiles in a large number of studies. To the extent that a single-determinant Kohn-Sham DFT approach is valid, so too will be the SEPDA electronic structure analysis technique.

As the reactants and products are identical in this case, only one side of the reaction coordinate was sampled. For the purposes of visual representation, we chose to define the deformation density of $E(\vec{R}_{yz})$ as the difference between the extracules of the transition state, D, and the reactant state, A. As the atoms within the system are moving throughout the course of the reaction, the extracule figures display overlaid structures of the reactant and transition states.

Our $E(\vec{R}_{yz})$ surfaces illustrate in precise detail how the centre-of-mass of the relevant electron pairs migrate i) toward the reaction centre in the case of the nucleophilic chloride lone pair (top), and ii) away from the reaction centre for the leaving group (bottom) as the reaction progresses from reactants (A) to transition state (D) (Figure 3.6). Additionally, the results observed from $P(u)$ afford a quantitative picture of the *relative* distribution of each Lewis pair. In the case of the chloride ion nucleophile, as the reaction progresses, the electrons do begin to separate to facilitate the interaction between the nucleophile and the electrophilic carbon centre. We can see this both in the shift of the maximum in $P(u)$ to larger u and in

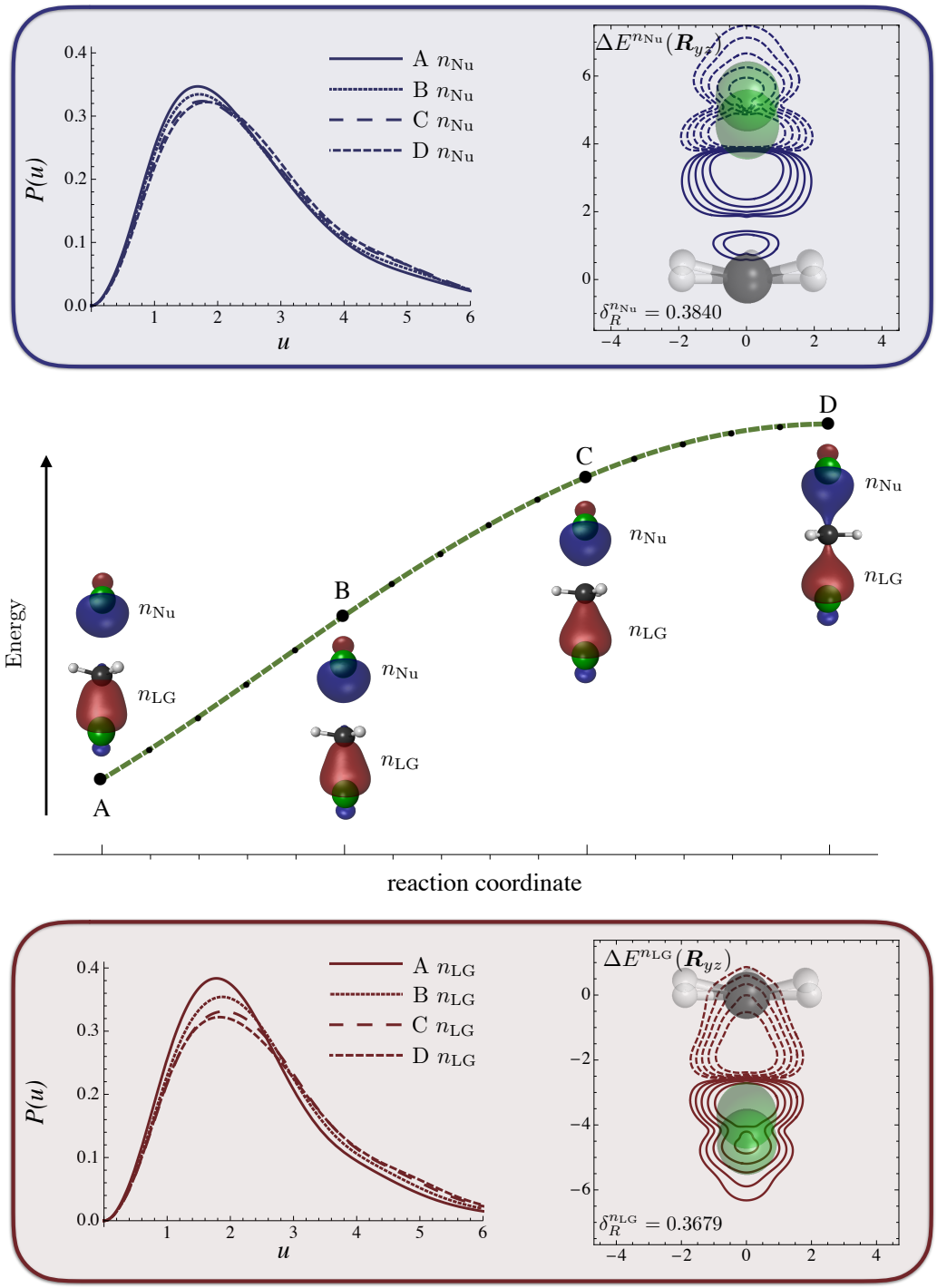


Figure 3.6: $P(u)$ and $\Delta E(\vec{R}_{yz})$ for the nucleophile and leaving group LMOs of an S_N2 reaction as it progresses from reactants (A) to transition state (D). The reaction profile depicts the four states (A-D) modeled for $P(u)$ along with depictions of the LMOs of the nucleophile and leaving group. Structures and energies were calculated at the OLYP/6-311G(d,p) level of theory based on the benchmark studies.[4] Solid and dashed lines in the $\Delta E(\vec{R}_{yz})$ plots denote positive and negative contours, respectively. Contours are plotted for $\pm n \times 10^{-3}$ where $n = 2, 4, 8, 16, 32$.

the concomitant broadening of the overall curve. The difference between each $P(u)$ curve is subtle however, which indicates that the formation of the covalent bond is more the result of the through-space migration of the localized pair of electrons, as opposed to the "spreading out" of the electron pair to form the bond.

In the case of the leaving group, the behaviour observed in $P(u)$ for the electron pair in the breaking bond is markedly different and not monotonic. Initially (A$\rightarrow$B), the interelectronic separation of the leaving group Lewis pair expands as is evidenced by the broader $P(u)$ distribution and the shifted maxima; subsequently however, the maxima in $P(u)$ recedes to smaller u despite a continually broadening distribution overall. This very detailed picture of the evolution of a bond breaking process is indicative of the accumulation of the electron pair at the chlorine centre (as the maximum in $P(u)$ recedes to smaller interelectronic separation) coinciding with the lengthening of the breaking bond (corresponding to a broader distribution overall).

We suspect that this very detailed, visual and intuitive analysis technique will be of broad utility in a wide variety of chemical contexts.

3.4 Conclusion

In the current chapter, we outlined a novel technique for predicting and interpreting chemical structure and behaviour by resorting to a fully quantum mechanical depiction of the familiar Lewis electron pair. By predicting a series of interelectronic distribution functions of individual pairs of electrons within an arbitrary chemical system we have shown that a wide range of chemical properties and phenomena

may be studied and interpreted using this technique. This Single Electron Pair Distribution Analysis (SEPDA) is a novel approach to answering key quantitative questions about the distribution of the well-known Lewis pairs, such as how they are distributed in space and how their relative velocities change in various chemical contexts.

We have shown that SEPDA may be used to quantify and classify myriad interactions including chemical bonding and non-covalent interactions. The nature of non-covalent interactions (as well as indications of their strength) may also be gleaned from such distributions and SEPDA can be used as an important tool to differentiate between interaction types.

Though we have chosen to focus our presentation on the so-called 'intuitive' chemical orbitals of Edmiston and Ruedenberg, it should be noted that the SEPDA technique may be readily applied to any orbital type.

While this chapter has focused on a couple of different examples of many different types of interactions with various electron pair densities, the remaining chapters will discuss specific applications in detail.

Chapter 4

Exploring Electron Pair Behaviour in Chemical Bonds Using the Extracule Density

This chapter has been reproduced with modifications with permission from Proud, A.J.; Mackenzie, D.E.C.K.; Pearson, J.K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 20194-20204. The majority of the work was done by Proud; however, the early coding of main.f90 was completed by Mackenzie.

4.1 Introduction

The molecular wavefunction contains a vast array of information; however, the Schrödinger equation[10] consists solely of one and two-electron operators. Thus, much of the information contained in the wavefunction is superfluous. As noted in Chapter 1, Hohenberg and Kohn[42] demonstrated that the energy of a chemical

system can be obtained, using only the electron density, $\rho(\mathbf{r})$,

$$\rho(\mathbf{r}) = \int |\Psi(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_N})|^2 \mathrm{d}\mathbf{r_2} \ldots \mathrm{d}\mathbf{r_N} \tag{4.1}$$

While one may be able to determine the energy from $\rho(\mathbf{r})$, extracting useful information regarding electron-electron interactions is inherently non-intuitive.

Instead, we can focus on the numerous electron pair densities that have been previously described. While the position intracule and intracules of a whole have been extensively studied, this is not as true of the extracule family. The main deficiency in $P(u)$ is the absence of any *absolute* position information. It provides no insight as to where in the molecular system the electrons are most likely to reside.[65] One way we can extract such information is through the extracule density, $E(\mathbf{R})$.[75,106] Recall that this density can be obtained from the pair density by:

$$E(\mathbf{R}) = \int \rho(\mathbf{r_1}, \mathbf{r_2}) \delta\left(\mathbf{R} - \tfrac{\mathbf{r_1}+\mathbf{r_2}}{2}\right) \mathrm{d}\mathbf{r_1} \mathrm{d}\mathbf{r_2} \tag{4.2}$$

The extracule density was first described by Coleman in the late 1960s[74]; however, it wasn't until the early 1980s when the first calculations of $E(\mathbf{R})$ were carried out.[75] Since the seminal paper by Thakkar and Moore, studies regarding the topology of $E(\mathbf{R})$ have largely been focused on the spherically averaged extracule density, $E(R)$, defined as

$$E(R) = \int E(\mathbf{R})\, \mathrm{d}\Omega_R \tag{4.3}$$

or a single dimension of $E(\mathbf{R})$.[100,117,118,142–148] While this scalar form may be useful

for linear systems, the complexity of the interpretation for large systems is readily apparent. Knowing only the distance of the centre-of-mass from a specified origin is not generally very informative in a three-dimensional molecule. For this reason, we have focused on the more topologically rich, $E(\mathbf{R})$ for the purposes of the study presented in this chapter.

While $E(\mathbf{R})$ does offer more clarity into the distribution of the centre-of-mass, difficulty still arises in its interpretation when one considers arbitrarily large 3D structures. SEPDA offers the ability to analyze a single pair of electrons. We demonstrate in this chapter the utility of this tool as we examine, in depth, the information present in $E(\boldsymbol{R})$.

4.2 Computational Methods

If the pair density is determined from a RHF wavefunction, equation (4.2) for a single molecular orbital may be expressed as

$$E(\mathbf{R}) = \sum_{\mu\nu\lambda\sigma}^{K} c_\mu c_\nu c_\lambda c_\sigma (\mu\nu\lambda\sigma)_E \tag{4.4}$$

where c_i describes the contribution of atomic orbital ϕ_μ to the molecular orbital of interest and $(\mu\nu\lambda\sigma)_E$ are the extracule integrals evaluated over the basis functions μ, ν, λ, and σ. These integrals are described by

$$(\mu\nu\lambda\sigma)_E = \int \phi_\mu^* (\mathbf{r})\, \phi_\nu (\mathbf{r})\, \phi_\lambda^* (2\mathbf{R} - \mathbf{r})\, \phi_\sigma (2\mathbf{R} - \mathbf{r})\, \mathrm{d}\mathbf{r} \tag{4.5}$$

Herein, ϕ_i denotes basis function i. Thakkar and Moore developed a series of formulae for the determination of the extracule integrals, which are summarized below.[75] These equations were utilized to calculate the necessary integrals for the evaluation of $E(\mathbf{R})$ for a specific LMO. For a basis set consisting of Gaussian-type orbitals, the basic integral over four s-type Gaussians is given by

$$(ssss)_E = \left(\frac{4\pi}{\zeta+\eta}\right)^{3/2} \exp\left[-\frac{\zeta\eta(2\mathbf{R}-\mathbf{P}-\mathbf{Q})^2}{\zeta+\eta}\right] \times \exp\left[-\frac{\alpha\beta(\mathbf{A}-\mathbf{B})^2}{\zeta} - \frac{\gamma\delta(\mathbf{C}-\mathbf{D})^2}{\eta}\right] \tag{4.6}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ define the centres on which the Gaussian primitives, with exponents α, β, γ, and δ, respectively, are centred. These exponents comprise $\zeta = \alpha + \beta$ and $\eta = \gamma + \delta$. The variables $\mathbf{P}$ and $\mathbf{Q}$ are defined as follows:

$$\mathbf{P} = \frac{\alpha\mathbf{A}+\beta\mathbf{B}}{\zeta} \qquad \mathbf{Q} = \frac{\gamma\mathbf{C}+\delta\mathbf{D}}{\eta} \tag{4.7}$$

For integrals containing orbitals of higher angular momenta, these new integrals can be determined by multiplying the basic integral by the angular factors, T_x, T_y, and T_z:

$$(\mu\nu\lambda\sigma)_E = (ssss)_E T_x T_y T_z \tag{4.8}$$

To define the angular factors (we will define the variables with respect to the angular factor in the x-axis, T_x, but these can easily be adapted to determine the y and z

directional angular factors by using the respective components for those directions), we must first introduce the following three variables:

$$g_k = \left(-\frac{\zeta+\eta}{4\zeta\eta}\right)^k \times \frac{1}{k!} \tag{4.9}$$

$$h_k = (P_x + Q_x - 2R_x)^k \times \frac{1}{k!} \tag{4.10}$$

$$s_k = \frac{1}{g_k}\sum_{j=0}^{k/2} g_j h_{k-2j} \tag{4.11}$$

where $0 \leq k \leq l_\mu + l_\nu + l_\lambda + l_\sigma$ in which l_i denotes the angular momentum of Gaussian primitive i, in the x-axis. Using these newly defined variables, T_x can be computed using

$$T_x = \sum_{i=0}^{l_\mu+l_\nu} \Lambda_i(l_\mu, l_\nu, P_x - A_x, P_x - B_x, \zeta) \times \sum_{j=0}^{l_\lambda+l_\sigma} s_{i+j}\Lambda_j(l_\lambda, l_\sigma, Q_x - C_x, Q_x - D_x, \eta) \tag{4.12}$$

where

$$\Lambda_j(l_1, l_2, a, b, c) = \sum_{k=0}^{(l_1+l_2-j)/2} f_{2k+j}(l_1, l_2, a, b)\frac{(2k+j)!}{(4c)^{k+j}k!\,j!} \tag{4.13}$$

Herein, f_j can be defined as the polynomial coefficients obtained from:

$$\sum_{j=0}^{l_1+l_2} f_j(l_1, l_2, a, b)x^j = (x+a)^{l_1}(x+b)^{l_2} \tag{4.14}$$

As the three angular factors are equal to unity for a set of four s-type Gaussians, equation (4.8) is a general formula that can be used in the evaluation of all extracule integrals.

To calculate the extracule densities, a Mura-Knowles grid[150] was utilized to determine the value of $E(R_x, R_y, R_z)$ at each grid point which was succeeded by interpolation to yield the required functions. This grid was adapted for the inclusion of negative values by incorporating grid points in both the positive and negative directions to ensure that all relevant spatial regions of the chemical system were adequately described. In order to obtain grids that were sufficiently dense to converge the resulting extracules, 201 points (100 points in each of the positive and negative directions in addition to the origin) were used in two of the three dimensions. These calculations would scale as $(K^4) \times (n_p^d)$ where d indicates the number of dimensions sampled and n_p defines the number of grid points in each dimension. As one dimension would necessarily be averaged through integration for purposes of visual representation, the benefit of sampling all three dimensions was deemed to be insufficient to warrant the computational cost. Thus, for the extracule density analysis of all molecular systems, the atoms important to the analysis were positioned in the yz-plane over which the grid was constructed.

All calculations were performed at the RHF/u6-311G(d,p) level of theory where u indicates that the basis set was completely uncontracted. HF calculations are adequate for these systems as the goal is to demonstrate the utility of the extracule density in the LPM as an interpretive tool in chemistry; however, should an alternative approach be desired, Kohn-Sham orbitals[43] could be employed in an identical

fashion. All geometry optimizations and LMO determinations were performed using the GAMESS software package.[20] Vibrational frequency analyses were conducted to ensure that the geometry represented an energy minimum. After computing the extracule densities using the Mura-Knowles grid, the data was analyzed using the Mathematica 8 software package.[156] Atomic units are used throughout.

4.3 Results and Discussion

4.3.1 Covalent Bonding

We begin with the hydrides of first row elements, i.e. LiH to HF. Two sets of analyses were carried out for this set. In Case 1, full geometry optimizations were performed on the molecules and these optimized structures were used in the analysis. For Case 2, the average bond length for the X-H (X = Li – F) bonds from Case 1 were determined and the optimized structures were then modified to include this average bond length *solely* for the bond to be analyzed. Thus, for NH_3, two of the N-H bonds would remain at the length determined through the geometry optimization, and only the one bond that was to be analyzed was adjusted to the average X-H bond length determined from the set of hydrides. Case 2 allows for a convenient comparison of the bond extracule densities as the nuclei involved in the hydride-bond LMO are at the same positions in Cartesian space.

Figure 4.1a) depicts the extracule density for the C-H bond in methane. One might expect that the maximum in the density would occur closer to the carbon atom considering its slightly more electronegative nature; however, one must con-

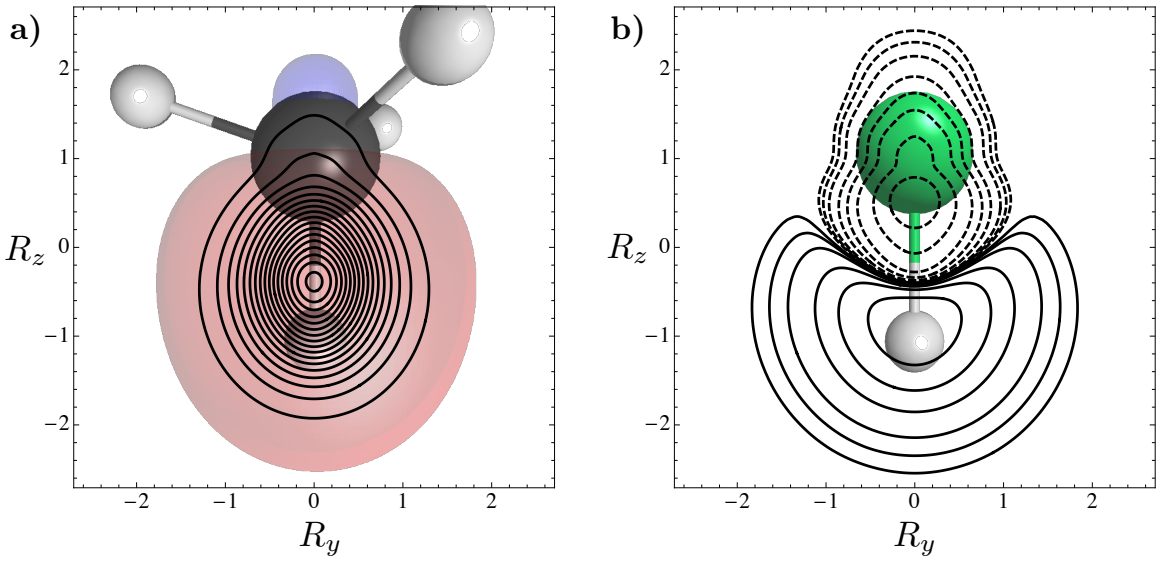


Figure 4.1: a) Depiction of $E(0, R_y, R_z)$ for the C-H bond in CH_4 with an overlay of the LMO for the bond and b) $\Delta E^{CH_3,F}(0, R_y, R_z)$ for the X-H bond LMO. Contour values were chosen as $m \times 10^{-n}$, where $m = 2, 4$ and 8 and $n = 3, 2,$ and 1 (the dashed lines signify negative contours).

sider the structure of the molecular orbital. As observed from the overlaid orbital representation, the C-H bond LMO largely resembles depictions of sp^3 orbitals common in freshman and organic chemistry textbooks.[90,157] Much of the density of the orbital extends out from the carbon atom beyond the hydrogen atom. This results in the electrons in the orbital being shifted more towards the hydrogen atom than one might initially expect.

What is more enlightening is observing the shift in the maximum as one changes the heavy atom from the highly electropositive Li atom to the highly electronegative F atom. These results are tabulated in Table 4.1. For all systems, the bond midpoint is positioned at (0,0,0) in Cartesian space. For the purposes of this discussion, the only coordinate listed is that which occurs along the bond axis, R_z. The R_y coordinate was close to 0 for all cases, but the small deviations can be explained by

the asymmetric inductive effects caused by the other atoms within the molecules. As the heavy atom in the molecule is modified from Li to F, the maximum in the extracule density shifts toward the heavy atom. What may be less obvious is the extent to which the centre-of-mass is shifting. This becomes far more evident as we examine $R_z^{\max}$ for Case 2. For all molecules in Case 2, the two nuclei are positioned at $R_z = \pm 1.08646$ with the heavy atom residing in the positive direction. For the case of the Li atom, the maximum in $E(0, R_y, R_z)$ occurs at $R_z = -1.028$, which is very close to the hydrogen nucleus. As we change our heavy atom from N to O, the centre-of-mass maximum shifts beyond the bond midpoint towards the heavy atom. In the most extreme case, HF, the maximum in the extracule density is found nearly halfway between the bond midpoint and the heavy F atom at $R_z = 0.500$. These same trends of the centre-of-mass shifting towards the electronegative atom are observed in the average R_z value, $\langle R_z \rangle$, which is given by

$$\langle R_z \rangle = \frac{1}{\langle R_{yz}^0 \rangle} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_z \times E(0, R_y, R_z) \mathrm{d}R_y \mathrm{d}R_z \tag{4.15}$$

where $\langle R_{yz}^0 \rangle$ is the zeroth moment in the bond plane (defined below in equation 4.16). As we consider only a slice of $E(\mathbf{R})$, $\langle R_z \rangle$ must be scaled by this value to obtain an accurate value for the average R_z. While the shift in the centre-of-mass towards the heavy atom as the electronegativity of that heavy atom increases is not surprising, it does demonstrate that the localized extracule density displays the effects one would expect based on chemical intuition. Less obvious is the trend

observed in $\langle R^0_{yz} \rangle$, which is defined as

$$\langle R^0_{yz} \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(0, R_y, R_z) \mathrm{d}R_y \mathrm{d}R_z \tag{4.16}$$

Because slices of the extracule density were chosen as opposed to averaging over one coordinate, the zeroth moment is not normalized to $\frac{N(N-1)}{2}$, as it otherwise would be. Instead, we obtain information regarding the amount of extricable density that resides within the given slice. Specifically, $\langle R^0_{yz} \rangle$, is the value of $E(R_x)$ where $R_x = 0$ and the remaining Cartesian coordinates have been averaged through integration. We note that $\langle R^0_{yz} \rangle$ is not bounded by 1 (as can be confirmed in Table 4.1), as a probability normally would be because it is not evaluated over a range in the x coordinate. One clear trend emerges as we change the identity of the heavy atom. The introduction of the more electronegative heavy atoms causes a contraction of the centre-of-mass to the bonding plane. While $\langle R^0_{yz} \rangle = 0.512$ in the case of LiH, this value nearly doubles to 0.982 upon replacing Li with the highly electronegative F. This observation is in strong agreement with the tendency of the electron-electron counterbalance density[158–162] for the helium isoelectronic series (from He to Ne^{8+}) to increase as the nuclear charge, and thus electronegativity, is increased.[160] Equivalent analyses were performed on the second row hydrides which are not shown as all of the trends were identical to those shown here.

The major benefit offered by the equidistant bond lengths analyzed in Case 2 is that one can accurately assess the extracule deformation density, $\Delta E^{\mathrm{X_1,X_2}}(0, R_y, R_z)$,

Table 4.1: Moments of $E(0, R_y, R_z)$ for the X-H bond LMO in first row hydrides.

	Case 1			**Case 2**		
System	$\langle R^0_{yz} \rangle$	$\langle R_z \rangle$	$R_z^{\max}$	$\langle R^0_{yz} \rangle$	$\langle R_z \rangle$	$R_z^{\max}$
LiH	0.473	-1.280	-1.453	0.512	-1.022	-1.028
BeH_2	0.557	-0.848	-0.996	0.595	-0.764	-0.836
BH_3	0.664	-0.579	-0.600	0.674	-0.570	-0.592
CH_4	0.743	-0.353	-0.383	0.728	-0.356	-0.385
NH_3	0.832	-0.181	-0.219	0.791	-0.157	-0.221
OH_2	0.932	-0.046	0.055	0.872	0.024	0.294
FH	1.047	0.091	0.272	0.982	0.218	0.500

which is given by

$$\Delta E^{\mathrm{X_1,X_2}}(0, R_y, R_z) = E^{\mathrm{X_1\text{-}H}}(0, R_y, R_z) - E^{\mathrm{X_2\text{-}H}}(0, R_y, R_z) \tag{4.17}$$

where X_i-H represents the LMO describing the hydride bond in the system of interest. With the positions of the two nuclei involved in the bond LMO fixed for all systems, all changes in $\Delta E^{\mathrm{X_1,X_2}}(0, R_y, R_z)$ can be attributed to the changing chemical environment. An example is shown in Figure 4.1(b) where X_1=CH_3 and X_2=F. As expected, the negative contours are present near the heavy atom indicating the greater presence of the electron pair centre-of-mass near the heavy atom in the HF system compared to CH_4. Likewise, positive values of $\Delta E^{\mathrm{CH_3,F}}(0, R_y, R_z)$ are present near the hydrogen atom as the electron density and consequently the centre-of-mass of the electron pair are drawn towards the fluorine atom in HF.

Following the analysis of the hydrides, the localized pair model was used to analyze compounds consisting of the -CH_3, -NH_2, -OH, and -F fragments from the first row and the -SiH_3, -PH_2, -SH, and -Cl fragments from the second row. Forming

covalent compounds from any two of these moieties results in 10 compounds from each of the first and second row and 16 compounds from the combination of building blocks from separate rows. Thus 36 compounds were constructed and analyzed in terms of the extracule densities for the bond between the two heavy atoms as well as the X-H bond for both heavy atoms in each fragment. As for the first and second row hydrides, the bond midpoint is positioned at (0,0,0) in Cartesian space with the heavy atom (or in the case of the X_1-X_2 bond, the X_1 atom) positioned in the positive R_z direction.

The results for these systems are listed in Table 4.2. For the LMO describing the bond between the two heavy atoms, X_1 and X_2, the same trends described in the previous section are evident for *most* species. We observe a substantial migration of the centre-of-mass into the bond plane as we increase the electronegativity of either heavy atom. Furthermore, the more obvious shifting of the centre-of-mass towards X_2 is evident as the electronegativity of X_2 increases. When considering a substitution from first row to second row heavy atoms, a significant decrease in $\langle R^0_{yz} \rangle$ is observed universally. This is indicative of a lower likelihood of observing $\boldsymbol{R}$ in the yz bond plane. Considering the significant size disparity between these rows, one might expect such a trend as the electrons in second row atoms would accommodate a larger volume outside of the selected bond plane leading to this observed decrease in $\langle R^0_{yz} \rangle$.

Once we shift to the X_i-H bond LMO, X_j is no longer directly part of the LMO of interest, but is instead separated by one bond. As LMOs are, by definition, *localized*, the changes in the extracule density are expected to be minimal when the two atoms

Table 4.2: Analysis of the X_1-X_2 and X_i-H bond LMOs in small first and second row compounds.

	X_1-X_2 Bond			X_1-H Bond			X_2-H Bond		
System	$\langle R^0_{yz}\rangle$	$\langle R_z\rangle$	$R_z^{\max}$	$\langle R^0_{yz}\rangle$	$\langle R_z\rangle$	$R_z^{\max}$	$\langle R^0_{yz}\rangle$	$\langle R_z\rangle$	$R_z^{\max}$
CH_3CH_3	0.752	0.000	0.000	0.749	-0.362	-0.386	0.749	-0.362	-0.386
CH_3NH_2	0.828	-0.183	-0.111	0.750	-0.384	-0.388	0.841	-0.189	-0.220
CH_3OH	0.928	-0.348	-0.493	0.759	-0.357	-0.370	0.950	-0.047	0.025
CH_3F	1.037	-0.477	-0.686	0.760	-0.365	-0.365	- - -	- - -	- - -
CH_3SiH_3	0.636	0.488	0.487	0.742	-0.335	-0.382	0.609	-0.865	-0.968
CH_3PH_2	0.661	0.298	0.284	0.749	-0.340	-0.374	0.651	-0.672	-0.771
CH_3SH	0.686	0.063	0.173	0.754	-0.331	-0.369	0.692	-0.465	-0.592
CH_3Cl	0.722	-0.154	-0.036	0.761	-0.333	-0.357	- - -	- - -	- - -
NH_2NH_2	0.885	0.004	0.000	0.843	-0.199	-0.217	0.843	-0.199	-0.217
NH_2OH	0.958	-0.174	-0.064	0.852	-0.190	-0.205	0.956	-0.057	0.030
NH_2F	1.054	-0.358	-0.630	0.858	-0.184	-0.195	- - -	- - -	- - -
NH_2SiH_3[a]	0.743	0.736	0.848	0.830	-0.209	-0.204	0.612	-0.888	-0.974
NH_2PH_2[a]	0.750	0.532	0.655	0.834	-0.181	-0.200	0.654	-0.702	-0.778
NH_2SH	0.764	0.333	0.323	0.842	-0.182	-0.199	0.695	-0.469	-0.583
NH_2Cl	0.775	0.102	0.188	0.854	-0.183	-0.194	- - -	- - -	- - -
OHOH	1.011	0.008	0.000	0.963	-0.044	0.060	0.963	-0.044	0.060
OHF	1.093	-0.197	-0.044	0.970	-0.036	0.080	- - -	- - -	- - -
$OHSiH_3$	0.872	0.791	0.893	0.939	-0.020	0.047	0.613	-0.881	-0.962
$OHPH_2$	0.872	0.678	0.843	0.946	-0.026	0.048	0.656	-0.700	-0.771
OHSH	0.863	0.532	0.798	0.954	-0.033	0.053	0.697	-0.473	-0.582
OHCl	0.857	0.343	0.283	0.960	-0.030	0.067	- - -	- - -	- - -
FF	1.152	0.000	0.000	- - -	- - -	- - -	- - -	- - -	- - -
$FSiH_3$	1.008	0.857	0.953	- - -	- - -	- - -	0.615	-0.860	-0.940
FPH_2	0.998	0.788	0.942	- - -	- - -	- - -	0.659	-0.678	-0.751
FSH	0.978	0.693	0.939	- - -	- - -	- - -	0.700	-0.458	-0.569
FCl	0.965	0.556	0.919	- - -	- - -	- - -	- - -	- - -	- - -
SiH_3SiH_3	0.502	0.000	0.000	0.608	-0.837	-0.962	0.608	-0.837	-0.962
SiH_3PH_2	0.544	-0.251	-0.033	0.611	-0.838	-0.952	0.647	-0.644	-0.765
SiH_3SH	0.594	-0.489	-0.458	0.613	-0.838	-0.943	0.687	-0.443	-0.589
SiH_3Cl	0.652	-0.673	-0.652	0.615	-0.829	-0.930	- - -	- - -	- - -
PH_2PH_2	0.574	0.002	0.000	0.652	-0.649	-0.757	0.652	-0.649	-0.757
PH_2SH	0.612	-0.274	-0.031	0.655	-0.652	-0.750	0.690	-0.447	-0.582
PH_2Cl	0.662	-0.519	-0.541	0.658	-0.644	-0.739	- - -	- - -	- - -
SHSH	0.642	-0.011	0.000	0.694	-0.448	-0.576	0.695	-0.313	-0.576
SHCl	0.678	0.270	0.024	0.698	-0.438	-0.566	- - -	- - -	- - -
ClCl	0.710	0.000	0.000	- - -	- - -	- - -	- - -	- - -	- - -

[a] The LMO for this X_1-X_2 bond showed significant distortion relative to the others. The maximum in $E(0, R_y, R_z)$ deviated from the bond axis, R_z, by > 0.050 a.u.

comprising the bond LMO remain the same. In these cases, the trends appear to vanish; however, there are a number of competing factors at play. First, we have

the aforementioned enhancement of the inductive effect caused by the increasing electronegativity of X_2 drawing the electron density, and thus the centre-of-mass towards X_2. Other related effects include the effect that X_2 has on the shape of the localized orbital in question as well as its effect on the bond lengths between X_1-H and X_1-X_2. Considering all of these factors, it is not surprising that there is no obvious trend in $\langle R_z^1 \rangle$ and R_z^{max} for the X-H bonds in these systems. However, the increase in $\langle R_{yz}^0 \rangle$ is still evident as the electronegativity of the non-participating heavy atom is increased.

To further explore the effects of electronegativity on the extracule density, we analyzed methane with varying levels of halogenation. Both the C-X and CH bonds of $CH_{4-n}F_n$ and $CH_{4-n}Cl_n$ (where $n = 0 - 4$) were explored. Figure 4.2 depicts the position of the maximum in the case of the C-H bond extracules for the $CH_{4-n}F_n$ systems. The introduction of the fluorine atoms cause an obvious shift in $\mathbf{R}_{\text{max}}$ within the C-H bond plane and *away* from the internuclear axis. The quantitative measures for these halogenated systems as well as those containing Cl are summarized in Table 4.3 and indicate that the extracular LPM has the capabilities to discern small but significant anisotropies in the topology of electron-electron interactions within the chemical bond.

With the overlaid structures depicting the positions of halogenation in Figure 4.2, the positioning is explained based on the inductive effects of the newly introduced electronegative atom. Upon the addition of the three halogens for the analysis of the C-H bond, the maximum returns to the bond axis due to the symmetry around the tetrahedral carbon, but it is significantly shifted towards the three

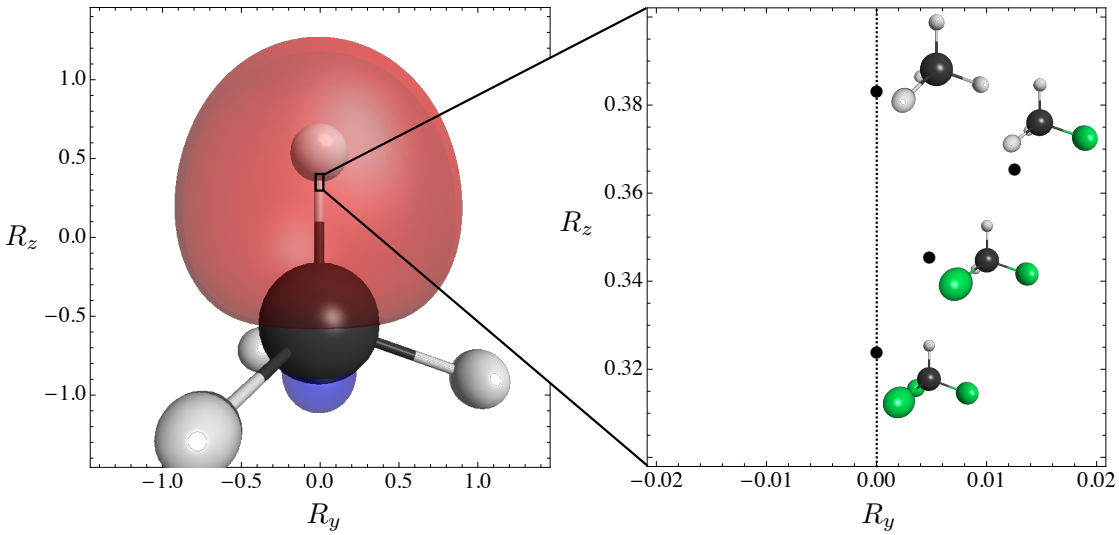


Figure 4.2: Pictorial representation of CH_4 to demonstrate the positions of each atom within the molecule combined with an inset of the positions of the maxima of $E(0, R_y, R_z)$ for the C-H bond in methane and its fluorinated derivatives, CH_nF_{3-n}, where $n = 1 - 3$ (the dashed line signifies the bond axis).

halogen atoms. These same trends can be observed for the C-H bonds in the chlorinated systems as well as the C-X bonds in both sets of halogenated molecules. Close analysis of the positions of the maxima reveals that the chlorine atoms tend to have a greater pull on the centre-of-mass in the C-H bonds than do the fluorine atoms. Due to the higher electronegativity of the fluorine species, this is rather surprising. The effect appears to be largest upon the introduction of the first chlorine atom. For the C-H bond, the change in R_z for CH_3F is -0.018 a.u. while the analogous value for CH_3Cl is -0.026 a.u. However, further halogenation does not lead to significant differences in the shift of the maxima between the fluorinated and chlorinated species. This same trend is seen for the C-X bond, but is less apparent due to the significant difference in the original position of the maxima. In fact for CHX_3 and CX_4 the shift in the maxima due to the substitution of an additional halogen is approximately three-fold greater in the case of X=F.

Table 4.3: Properties of $E(0, R_y, R_z)$ for halogenated derivatives of methane.

	C-H Bond		C-X Bond	
System	$\langle R^0_{yz} \rangle$	(R_y, R_z) of Max	$\langle R^0_{yz} \rangle$	(R_y, R_z) of Max
CH_4	0.745	(0.000, 0.383)	- - -	- - -
CH_3F	0.760	(0.013, 0.365)	1.037	(0.000, 0.686)
CH_2F_2	0.782	(0.005, 0.346)	1.053	(0.009, 0.651)
CHF_3	0.797	(0.000, 0.324)	1.065	(0.004, 0.618)
CF_4	- - -	- - -	1.070	(0.000, 0.589)
CH_4	0.745	(0.000, 0.383)	- - -	- - -
CH_3Cl	0.760	(0.015, 0.357)	0.722	(0.000, 0.036)
CH_2Cl_2	0.772	(0.007, 0.337)	0.732	(0.016, -0.116)
$CHCl_3$	0.783	(0.000, 0.320)	0.736	(0.007, -0.127)
CCl_4	- - -	- - -	0.743	(0.000, -0.136)

4.3.2 Bond Strain

An ideal tetrahedral carbon has sp^3 hybridization with bond angles of 109.5°. However, for some cycloalkanes, this conformation is simply not possible. For instance, cyclopropane, a well documented example,[163–165] contains significant amounts of strain due to its triangular conformation deviating significantly from the optimal configuration around a *tetrahedral* carbon. This strain causes the formation of "bent bonds" or "banana bonds". This bending is clearly observed in the extracule density of not only cyclopropane but also to some extent, in cyclobutane (Figure 4.3). The bending in $E(\mathbf{R})$ effectively vanishes in the densities calculated for cyclopentane and cyclohexane. A quantitative analysis can be performed by determining the position of the maximum in $E(\mathbf{R})$ outside of the bond axis, i.e. R_y. This measure is tabulated in Table 4.4. The maximum for cyclopropane is observed at $R_y^{\text{max}} = 0.349$ but migrates towards the bond axis as the ring strain decreases and essentially resides in the bond axis for the "strainless" cyclohexane.

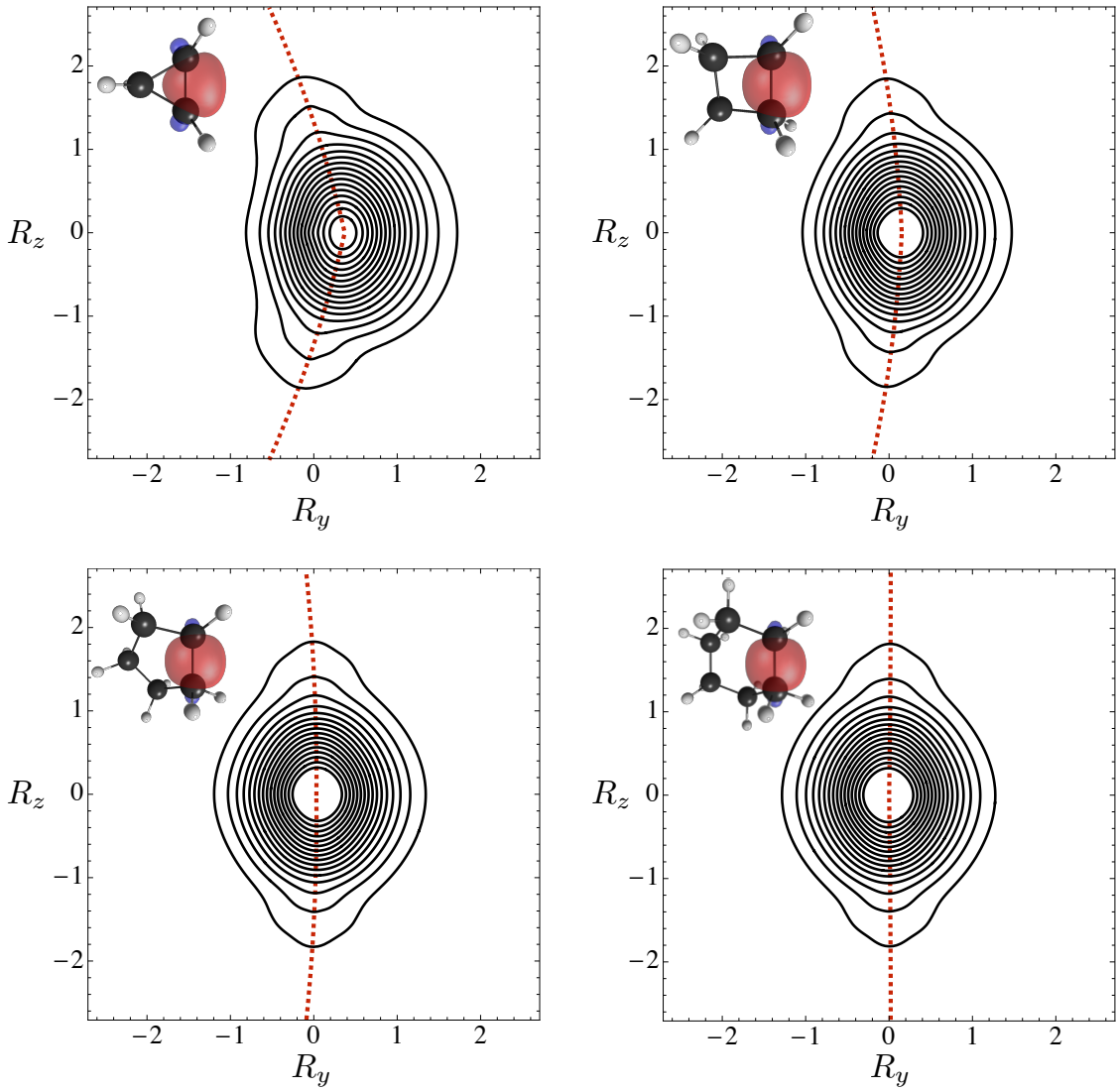


Figure 4.3: Depiction of $E(\mathbf{R})$ for representative C-C bonds in the cyclic systems ranging from cyclopropane to cyclohexane. Models of the appropriate molecule are inlayed in the top left hand corner of each graph to provide the reader with insight as to the spatial orientation of each molecule. The dashed line traces the curve of slowest descent in $E(R_y, R_z)$ to illustrate the deviation from the bond axis. Contour values were chosen as $0.02 \times n$ where $n = 1 - 16$.

To accommodate the smaller angles involving the C-C bonds in the smaller cycloalkanes, the orbitals take on significantly more p-character than a typical C-C bond between *tetrahedral* carbons. In the case of cyclopropane, the carbon atoms participating in the C-C bonds are considered to be sp^5 hybridized with respect to

Table 4.4: Properties of $E(0, R_y, R_z)$ for the C-C bonds in cycloalkanes.

System	$\langle R^0_{yz} \rangle$	$R_y^{\max}$
Cyclopropane (C_3H_6)	0.750	0.349
Cyclobutane (C_4H_8)	0.751	0.099
Cyclopentane (C_5H_{10})	0.755	0.017
Cyclohexane (C_6H_{12})	0.757	-0.004

that orbital.[163] As the size of the ring in the cycloalkane increases, the strain is relieved and the amount of p-character in the bonds decrease. As p-orbitals are less electronegative than s-orbitals, progressing from cyclopropane to cyclohexane, we would expect the electronegativity of the carbon involved in the C-C bond to increase, leading to an increase in the proportion of $E(\mathbf{R})$ present in the bond plane. This prediction is confirmed by the values of $\langle R^0_{yz} \rangle$ provided in Table 4.4.

4.3.3 Non-Covalent Interactions

While LMOs are largely local in nature, as demonstrated above, they are influenced in characteristic ways by their neighbouring chemical environments. This suggests that the LPM has utility in analyzing non-covalent interactions. For example, hydrogen bonding may be interpreted as the interaction between an electron rich lone pair of a donor species with an electron deficient acceptor species. The extent and/or character of the interaction may then be probed by observing changes in the distributions of localized electron pairs on either the donor or acceptor species (or both). Here, we have modelled the hydrogen bonding interaction between HF and $MeNH_2$ through the σ_{HF} bond LMO of HF and the n_N lone pair LMO on the $MeNH_2$ nitrogen. Accurate geometries for the hydrogen bonding complex were obtained

from the S66x8 data set.[151] Extracule calculations were performed on the H-F bond in the absence of $MeNH_2$ (and vice versa) and with varying separations between the HF and $MeNH_2$ molecules. The b_0 separation indicates that the distance between the hydrogen bond donor and acceptor is that which is obtained from the geometry optimization carried out at the MP2/cc-pVTZ level. Systems denoted by $x \times b_0$ indicate that the distance between the two species, d, is scaled proportionally to x. Thus, the HF-MeOH complex where $d = 2.0b_0$ contains an H-bond distance that is twice the value obtained in the geometry optimization. All other geometrical parameters remain the same. For the HF bond, the bond midpoint was positioned at the origin in Cartesian space, while for the lone pair, the nitrogen atom in methylamine was positioned at (0,0,0.945) while the H in hydrogen fluoride was positioned along the R_z axis at positions relative to the separation of the two species. The positioning of the nitrogen atom was chosen to allow for adequate sampling of $E(0, R_y, R_z)$ using the previously described grid points.

Our goal in analyzing the extracule density of these LMOs was to observe the effect on the extracule density as the hydrogen bond formed and how those effects varied as the distance between the two species grew. Thus we analyzed the extracule deformation density of orbital ϕ, $\Delta E_d^{\phi}(\mathbf{R})$, which we define in this case as

$$\Delta E_d^{\phi}(\mathbf{R}) = E_d^{\phi,\text{complex}}(\mathbf{R}) - E^{\phi,\text{molecule}}(\mathbf{R}) \tag{4.18}$$

which parametrically depends on d, the distance between the species in the molecular complex.

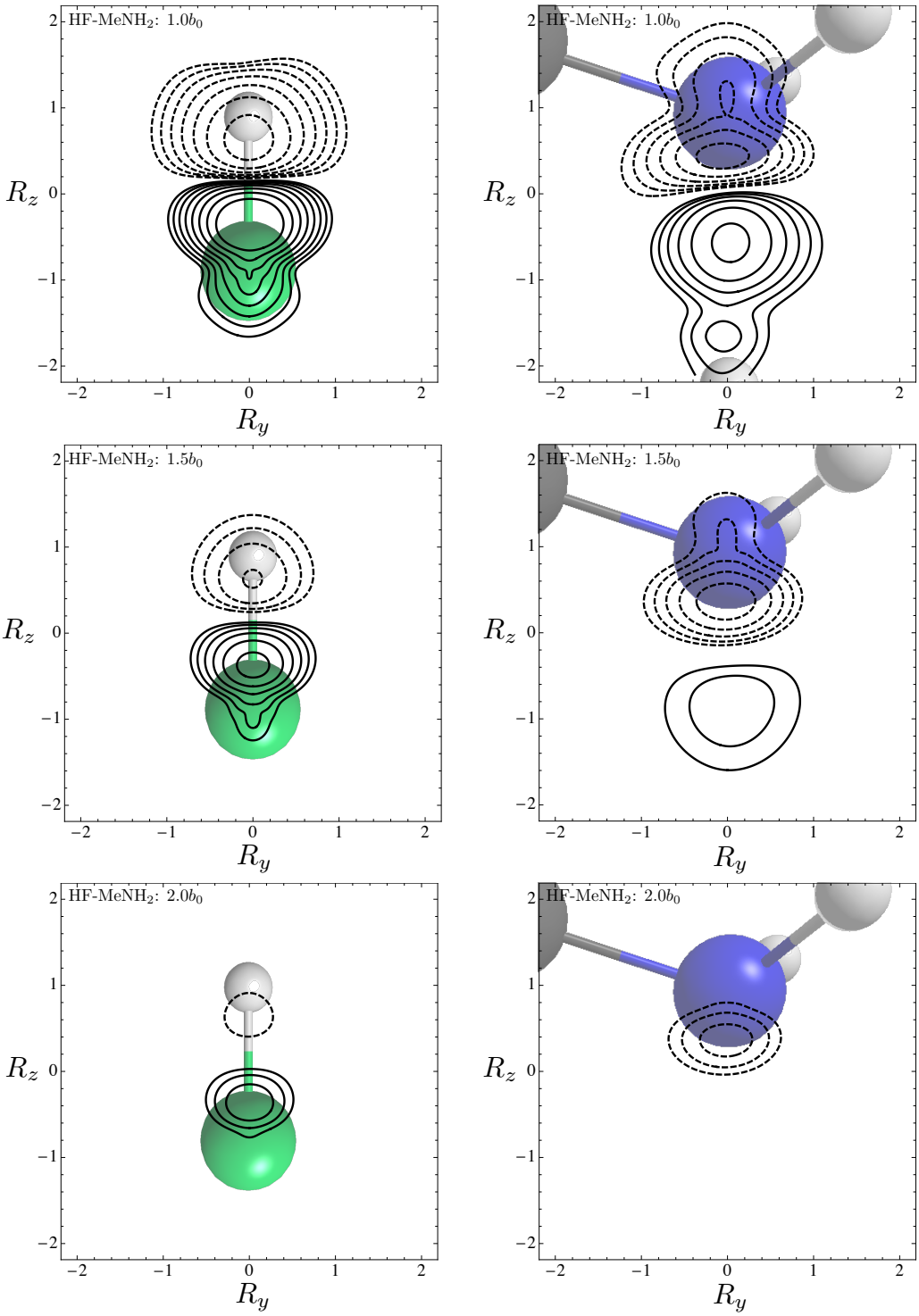


Figure 4.4: Depiction of $\Delta E_d^{\phi}(\mathbf{R})$ for the σ_{HF} bond LMO in H-F (left) and the n_{N} lone pair LMO in $MeNH_2$ (right) for the HF-$MeNH_2$ hydrogen bonded complex at various distances of separation, $x \times b_0$, between the donor and acceptor. Contours were chosen as $\pm 0.003 \times 1.5^n$ where $n = 1 - 8$. Negative contours are denoted by dashed lines.

These results are depicted in Figure 4.4. For the HF bond, the figure clearly shows an increase in $E(\boldsymbol{R})$ near the F atom as the hydrogen bond forms. This effect diminishes significantly as the distance separating the species varies from $d = 0.9b_0$ to $d = 2.0b_0$ ($0.9b_0$ not shown). This change can be concisely rationalized by the migration of electrons during the formation of a hydrogen bond. In this case, the H atom in H-F would interact with the electron density of the donor lone pair in $MeNH_2$. Interacting with the nitrogen allows the electrons within the H-F bond to migrate towards the F atom resulting in the increase in the likelihood of the centre-of-mass of that electron pair to be close to fluorine. Conversely, when considering the lone pair in $MeNH_2$, we observe a decrease in the extracule deformation density near the nitrogen atom and an increase in the internuclear region between N and the HF molecule. Unlike the H-F bond where the hydrogen-fluorine interaction was weakening, here the nitrogen-hydrogen interaction is becoming stronger. Thus, the electrons are migrating toward the hydrogen and consequently shifting the centre-of-mass away from the nitrogen atom resulting in the observed depletions in the extracule density in this area. This observed migration of electrons from the hydrogen to the fluorine in the H-F bond combined with the donation of electrons from nitrogen to the electron deficient hydrogen is in excellent agreement with the resonance-covalency[88,89] interpretation of hydrogen bonding (or any non-covalent interaction) as opposed to the more traditional dipole-dipole interaction interpretation.

To *quantify* the differences between the extracule densities of σ_{HF} and n_{N} before and after complexation, we have employed similar measures as noted previously in-

Table 4.5: Properties of $E_d^{\phi,\text{HF-MeNH}_2}(0, R_y, R_z)$ and $\Delta E_d^{\phi}(0, R_y, R_z)$ for the HF bond (σ_{HF}) and the $MeNH_2$ lone pair (n_{N}) LMOs.

	$E_d^{\phi,\text{HF-MeNH}_2}$		ΔE_d^{ϕ}	
Molecule, ϕ	$\langle R_{yz}^0 \rangle$	$R_z^{\max}$	δ_{yz}	$R_z^{\max}$
HF ($d = 0.9b_0$), σ_{HF}	1.074	-0.305	0.148	-0.386
HF ($d = 1.0b_0$), σ_{HF}	1.067	-0.298	0.108	-0.384
HF ($d = 1.5b_0$), σ_{HF}	1.051	-0.283	0.027	-0.381
HF ($d = 2.0b_0$), σ_{HF}	1.047	-0.279	0.010	-0.381
HF (no complex), σ_{HF}	1.045	-0.277	0.000	- - -
$MeNH_2$ ($d = 0.9b_0$), n_{N}	0.771	0.241	0.080	-0.439
$MeNH_2$ ($d = 1.0b_0$), n_{N}	0.759	0.248	0.067	-0.566
$MeNH_2$ ($d = 1.5b_0$), n_{N}	0.751	0.262	0.035	-0.837
$MeNH_2$ ($d = 2.0b_0$), n_{N}	0.755	0.265	0.013	-0.756
$MeNH_2$ (no complex), n_{N}	0.759	0.268	0.000	- - -

cluding the zeroth moment $\langle R_{yz}^0 \rangle$ of $E_d^{\sigma_{\text{HF}},\text{HF-MeNH}_2}(0, R_y, R_z)$ and $E_d^{n_{\text{N}},\text{HF-MeNH}_2}(0, R_y, R_z)$ as well as $R_z^{\max}$ of both the extracules and the extracule deformation densities. For $\Delta E_d^{\phi}(0, R_y, R_z)$, we also define a new measure, δ_{yz}, referring to the magnitude of the difference between the extracules:

$$\delta_{yz} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \Delta E_d^{\phi}(0, R_y, R_z) \right| \, \mathrm{d}R_y \, \mathrm{d}R_z \tag{4.19}$$

These metrics are all listed in Table 4.5. As before, the introduction of an electronegative species (N in $MeNH_2$) caused an increase in $\langle R_{yz}^0 \rangle$ for the HF bond LMO. This effect is even present in the case where $d = 2.0b_0$; it is small, but still significant. However, no trend is apparent in the zeroth moment for the lone pair. When observing the position of the maxima, as noted in the discussion of Figure 4.4, the maxima shift toward the F atom for the HF bond, while they shift away from the nitrogen atom, towards the acceptor species (HF) in the case of the $MeNH_2$ lone pair.

δ_{yz} provides us with an absolute measure of the variation in the extracule density for each LMO as a function of the intermolecular interaction. It is relatively large when $d = 0.9b_0$ ($\delta_{yz} = 0.148$ and $\delta_{yz} = 0.080$ for σ_{HF} and n_{N}, respectively) and decreases as the methylamine and hydrogen fluoride are separated. At $d = 2.0b_0$, δ_{yz} reduces to 0.010 (σ_{HF}) and 0.013 (n_{N}) suggesting that the strength of the hydrogen bonding interaction is related to δ_{yz}. Observing these changes in δ_{yz} can provide an indication of the strength of the interaction between the donor and acceptor species, especially when weighted against the energetic cost of nuclear repulsion with decreasing d. Further work to elucidate relationships between intermolecular interaction energies and electron pair distributions (intracular and/or extracular) in position and momentum spaces is ongoing in our laboratory.

4.4 Conclusions

Herein, we have introduced a novel tool for the analysis of electronic structure. While the extracule density has been studied in the past, the breadth of systems studied has been very limited. This could be due in part to the complexity involved in interpreting a probability density for $N(N-1)/2$ pairs of electrons. By accessing localized regions of chemical space through the use of ER localized molecular orbitals, we not only simplify the interpretation of the extracule density, but also afford a quantum mechanical interpretation of “chemically intuitive" features of electronic structure.

While this study only involved calculations performed at the HF level of theory, the general trends in chemical behaviour observed are not expected to change

through the use of correlated models. Regardless, the localized pair model does offer the capability to perform analyses using Kohn-Sham orbitals to account for correlation.[43] Studies are presently under way in our lab detailing the effects of correlation within localized chemical bonds for intracule densities and could easily be implemented for the study of extracule densities.

This study has demonstrated the types of information that can be extracted from the localized extracule density for simple systems, but one can extend these calculations to larger systems. The main obstacle to the study of large chemical systems is the time required for such calculations. However, through the use of LMOs, this barrier can be partially overcome by the inclusion of only atomic orbitals in close proximity to the molecular orbital under scrutiny. While this study was conducted with ER LMOs, one can apply the technique in an identical fashion to other localized orbitals, such as the previously mentioned NBOs, IBOs, and ALMOs, as well as any canonical molecular orbital of interest.

Chapter 5

Developing a Theoretical Model for Quantifying Electronegativity based on the Position Extracule

This chapter has been reproduced (plus significant additions) with permission from Proud, A.J.; Pearson, J.K. *Can. J. Chem.* **2016**, *94*, 1077-1081.

5.1 Introduction

Electronegativity is a ubiquitous chemical concept that is used to explain many periodic trends from polarity and partial charges to atomic size. Additionally, an intuitive grasp of electronegativity is paramount for the understanding of electrophilic/nucleophilic regions which help guide synthetic procedures, among many other important applications in chemistry. Interestingly, an actual *definition* of electronegativity has historically been somewhat elusive and there remains an appetite

for a simple, yet theoretically rigorous definition based on quantum mechanical properties.[166] The development of the concept of electronegativity is commonly attributed to Pauling; however, the idea is a perennial favourite among the chemical literature and has been discussed since the late 18th century.[167,168] Regardless, Pauling was the first to *quantify* the property and his electronegativity scale is still in common use today.[157,169] Pauling's definition of electronegativity, χ, was based on differences in bond dissociation energies (E_d) of two homonuclear diatomics, A−A and B−B, compared to that of the heteronuclear diatomic, A-B. This was expressed mathematically as

$$\chi_A - \chi_B = \frac{1}{\sqrt{eV}}\sqrt{E_d^{AB} - \frac{1}{2}[E_d^{AA} + E_d^{BB}]} \tag{5.1}$$

where the dissociation energies are expressed in terms of electron volts (eV). Since this initial quantification of electronegativity, there have been numerous different scales developed in an attempt to accurately quantify this highly useful property.[170–180] These scales are based on various physical properties including bond force constants[172], effective nuclear charge/covalent radii[171,175], ionization potentials/electron affinities[170], etc. Aside from these empirical approaches, Simons et al. developed the first purely theoretical approach to quantifying electronegativity.[181] This model was based on the positions of floating spherical Gaussian orbitals. However, Boyd and coworkers were the first to develop a theory for quantifying electronegativity based on the *topology* of the electron density, $\rho(\mathbf{r})$, despite what could be considered an intuitive relationship between the two.[182,183] Their work re-

vealed a power series relationship between the Pauling electronegativity of atom or group A (when bonded to H, thereby forming the system A−H) and a parameter, the electronegativity factor, F_A, which they define as

$$F_A = \frac{r_H}{N_A \rho(\mathbf{r_c}) r_{AH}} \tag{5.2}$$

Herein, $\mathbf{r_c}$ is the position of bond critical point of the A−H bond, r_{AH} is the bond length, r_H is the distance between the bond critical point and the hydrogen atom, and N_A is the number of valence electrons associated with atom/group A. While this form of analysis worked quite well for both atomic and group electronegativities, it is a somewhat convoluted relationship, especially considering the deceptive simplicity of the electronegativity concept itself. A more simple and intuitive approach would be ideal. Boyd and others recently commented on how a topological approach would be highly useful in the determination of various properties, one of which, was electronegativity. Boyd specifically stated the need for a more rigorous theoretical basis for electronegativity, noting that the most likely avenue for this basis would be through a topological approach.[166]

One way to approach this is to examine the topology of localized electron *pair* distributions.[82,83,91] The tools available in SEPDA should represent a suitable technique for the analysis of the topology of localized electron pairs in a chemical system from a purely theoretical standpoint. The LPM was developed specifically to analyze electron pair behaviour within covalent bonds and lone pairs (though it is generally applicable) and we are now beginning to demonstrate the wide range of

applications of such a technique.

In considering a localized electron pair approach, the most obvious way to observe the migration of an electron pair as a result of electronegativity differences is through the extracule density, $E(R)$. This way, we can monitor the change in the centre-of-mass of the electron pair from an arbitrarily chosen origin (in the current work, this origin corresponds to the position of the heavy-atom nucleus of the A$-$H bond). Recall that this can be obtained from the LMO of interest, ψ, by

$$E(R) = \langle\psi|\delta(R - \tfrac{|\mathbf{r_1}+\mathbf{r_2}|}{2})|\psi\rangle \tag{5.3}$$

These localized extracules can model the change in position of the electron pair centre-of-mass as a function of the two atoms involved in the bond, thereby providing a quantitative measure of the tendency for each atom to attract a *specific* electron pair to itself.

While monitoring the absolute position of the electron pair through the position extracule may be a more intuitive approach to modelling electronegativity, analyzing the relative positions of the electron pair through the position intracule could prove fruitful as well. Much like the extracule, the position intracule for the A$-$H bond can be determined from

$$P(u) = \langle\psi|\delta(u - |\mathbf{r_1} - \mathbf{r_2}|)|\psi\rangle \tag{5.4}$$

The main focus of the study presented in this chapter is on the relationship between

electronegativity and the extracule density; however, the intracular analysis is also provided to demonstrate that while less intuitively related to electronegativity, it can yield significant insight as well.

For the purposes of this study, hydrogen was bonded to the atom of interest (as it is considered to be neither electron donating, nor electron withdrawing), thereby forming the system A−H, and representing a unique way of quantifying the electronegativity of atom A, based purely on topological properties of electron pairs. The systems of interest in this study are those where X is any first or second row atom, excluding noble gases.

While the main goal is not to exactly match existing electronegativity scales, considering that most existing scales agree well with one another, strong correlations to existing models would provide evidence for the validity of this model. Perfect agreement with existing models is not necessary however, as existing models do not even perfectly agree with one another. Considering so many different models have been developed, this points to the absence of a fundamentally strong definition of electronegativity. The model presented herein represents a very simple and intuitive model of electronegativity as it explores the purely quantum mechanical topology of electron *pair* distributions.

5.2 Computational Methods

The extracule densities for the systems involved in this study were calculated using a modified version of the recurrence relation developed by Hollett and Gill for position intracules[73]. Before we can describe the recurrence relation, we must first

introduce a few variables. The recurrence relation is developed for a set of Gaussian functions where a Gaussian primitive, **a**, with exponent α, is given by

$$|\mathbf{a}\rangle = (x - A_x)^{a_x}(y - A_y)^{a_y}(z - A_z)^{a_z} e^{-\alpha|\mathbf{r}-\mathbf{A}|^2} \tag{5.5}$$

where the angular momentum of $\mathbf{a} = (a_x, a_y, a_z)$ and the function is centred on $\mathbf{A} = (A_x, A_y, A_z)$. Likewise, the Gaussian primitives, $|\mathbf{b}\rangle$, $|\mathbf{c}\rangle$, and $|\mathbf{d}\rangle$ are centred at $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ with exponents β, γ, and δ, respectively. Using these Gaussian type orbitals, we can define the following variables:

$$\nu^2 = \frac{(\alpha+\beta)(\gamma+\delta)}{\alpha+\beta+\gamma+\delta} \qquad S_{ab} = \text{Exp}\left[-\frac{\alpha\beta|\mathbf{A}-\mathbf{B}|^2}{\alpha+\beta} - \frac{\gamma\delta|\mathbf{C}-\mathbf{D}|^2}{\gamma+\delta}\right] \tag{5.6}$$

$$\mathbf{U}^E = \frac{\alpha\mathbf{A}+\beta\mathbf{B}}{\alpha+\beta} + \frac{\gamma\mathbf{C}+\delta\mathbf{D}}{\gamma+\delta} \qquad \mathbf{U}^P = \frac{\alpha\mathbf{A}+\beta\mathbf{B}}{\alpha+\beta} - \frac{\gamma\mathbf{C}+\delta\mathbf{D}}{\gamma+\delta} \tag{5.7}$$

The recurrence relation for both the position extracule is given by the following 8-term recursive formula:

$$\begin{aligned}
[(\mathbf{a}+\mathbf{1_i})\mathbf{bcd}]^{(l)} = {} & \frac{\beta(B_i - A_i)}{\alpha+\beta}[\mathbf{abcd}]^{(l)} + \frac{U_i^E}{\alpha+\beta}[\mathbf{abcd}]^{(l+1)} \\
& + \frac{a_i}{2(\alpha+\beta)}[(\mathbf{a}-\mathbf{1_i})\mathbf{bcd}]^{(l)} + \frac{a_i}{2(\alpha+\beta)^2}[(\mathbf{a}-\mathbf{1_i})\mathbf{bcd}]^{(l+1)} \\
& + \frac{b_i}{2(\alpha+\beta)}[\mathbf{a}(\mathbf{b}-\mathbf{1_i})\mathbf{cd}]^{(l)} + \frac{b_i}{2(\alpha+\beta)^2}[\mathbf{a}(\mathbf{b}-\mathbf{1_i})\mathbf{cd}]^{(l+1)} \\
& + \frac{c_i}{2(\alpha+\beta)(\gamma+\delta)}[\mathbf{ab}(\mathbf{c}-\mathbf{1_i})\mathbf{d}]^{(l+1)} \\
& + \frac{d_i}{2(\alpha+\beta)(\gamma+\delta)}[\mathbf{abc}(\mathbf{d}-\mathbf{1_i})]^{(l+1)}
\end{aligned} \tag{5.8}$$

while that for the position intracule is:

$$
\begin{aligned}
[(\mathbf{a}+\mathbf{1_i})\mathbf{bcd}]^{(l)} = {} & \frac{\beta(B_i - A_i)}{\alpha+\beta}[\mathbf{abcd}]^{(l)} + \frac{U_i^P}{\alpha+\beta}[\mathbf{abcd}]^{(l+1)} \\
& + \frac{a_i}{2(\alpha+\beta)}[(\mathbf{a}-\mathbf{1_i})\mathbf{bcd}]^{(l)} + \frac{a_i}{2(\alpha+\beta)^2}[(\mathbf{a}-\mathbf{1_i})\mathbf{bcd}]^{(l+1)} \\
& + \frac{b_i}{2(\alpha+\beta)}[\mathbf{a}(\mathbf{b}-\mathbf{1_i})\mathbf{cd}]^{(l)} + \frac{b_i}{2(\alpha+\beta)^2}[\mathbf{a}(\mathbf{b}-\mathbf{1_i})\mathbf{cd}]^{(l+1)} \\
& - \frac{c_i}{2(\alpha+\beta)(\gamma+\delta)}[\mathbf{ab}(\mathbf{c}-\mathbf{1_i})\mathbf{d}]^{(l+1)} \\
& - \frac{d_i}{2(\alpha+\beta)(\gamma+\delta)}[\mathbf{abc}(\mathbf{d}-\mathbf{1_i})]^{(l+1)}
\end{aligned}
\tag{5.9}
$$

These recurrence relations describe how to obtain the integral for augmenting the angular momentum of the i^{th} coordinate ($i = x$,y, or z) by one ($\mathbf{1_i}$). In order to determine the required integral $[(\boldsymbol{a}+\mathbf{1}_{\boldsymbol{i}})\mathbf{bcd}]^{(l)}$, one needs the fundamental integral $[\mathbf{0000}]^l$ which will be outlined for each electron pair density in the next section. The derivation of this recurrence relation is provided in detail in Appendix B.

5.2.1 Extracule and Intracule Analysis

The fundamental integral for the scalar position extracule density is given by

$$
[\mathbf{0000}]^{(l)} = \frac{32\pi^{5/2}R^2 e^{-4\nu^2R^2}S_{ab}}{(\alpha+\beta+\gamma+\delta)^{3/2}}\left(\frac{\partial}{\partial \mathbf{U}^2}\right)^l [e^{-\nu^2U^2} i_0(4\nu^2UR)] \tag{5.10}
$$

Similarly, for the position intracule density, the fundamental integral is

$$
[\mathbf{0000}]^{(l)} = \frac{4\pi^{5/2}u^2 e^{-\nu^2u^2}S_{ab}}{(\alpha+\beta+\gamma+\delta)^{3/2}}\left(\frac{\partial}{\partial \mathbf{U}^2}\right)^l [e^{-\nu^2U^2} i_0(2\nu^2Uu)] \tag{5.11}
$$

where $i_0(x)$ is the modified spherical Bessel function of the first kind defined as $i_0(x) = \sinh(x)/x$.

Two quantitative metrics describing the location of the electron pair centre-of-mass, R, were chosen for the purposes of this study. They are the position of the maximum of the extracule density, $R^{\max}$, and the first moment, or average value of R for a particular $E(R)$, $\langle R \rangle$, which is defined as

$$\langle R \rangle = \int_0^\infty R \times E(R) \mathrm{d}R \tag{5.12}$$

As defined, these metrics can be problematic as they are bond length dependent. Thus, for comparative purposes, they were determined as a fraction of the bond length and are denoted $R_{\%}^{\max}$ and $\langle R \rangle_{\%}$. Since both metrics are determined as a ratio with respect to the bond length, all quantities expressed throughout are unitless.

Similarly, for $P(u)$, three metrics were explored. The two intracular analogues of the extracular metrics were chosen, i.e. the position of the maximum of the intracule density, $u^{\max}$, and the average value of u, given by $\langle u \rangle$. The intracule also possesses another easily obtained metric with great significance. The first inverse moment of the intracule density is given by

$$\langle u^{-1} \rangle = \int_0^\infty \frac{1}{u} P(u) \mathrm{d}u \tag{5.13}$$

which is equivalent to the electron repulsion energy in the system. Due to the physical significance of the moment, it was included in the analysis as well.

For the first part of this study, all chemical systems were optimized at the Restricted open-shell Hartree-Fock level of theory (ROHF for radicals) with a u6-311G(d,p) basis set where the u indicates that the basis set was completely uncontracted. In the second part of the study, various density functionals are assessed to determine whether the agreement to existing electronegativity scales improve upon the introduction of electron correlation energy.

5.3 Results and Discussion

5.3.1 Hartree-Fock Method

Extracule densities were determined for the A−H bond in each of the saturated first and second row hydrides. Each geometry was optimized at the same level of theory and the extracule was determined for the bond LMO describing said bond. The extracules for these hydrides are shown in Figures 5.1 a) and 5.1 b). As expected, as we move across the periodic table from left to right, the electron pair centre-of-mass migrates from the H atom (i.e. $R = 100\%$) towards the heavy atom (i.e.

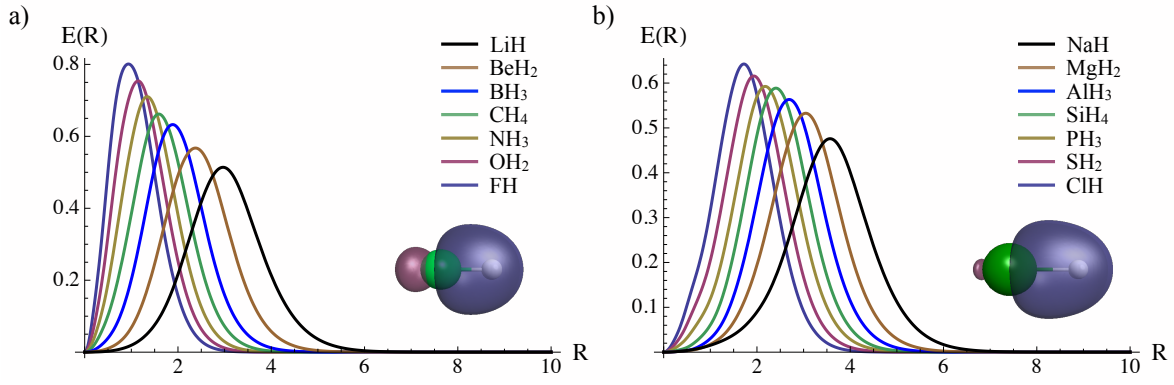


Figure 5.1: Localized extracule densities for the A−H bond in saturated a) first row hydrides, and b) second row hydrides with insets of the A−H bond LMO of F−H and Cl−H for illustrative purposes.

$R = 0\%$) due to the increase in electronegativity. This is clearly demonstrated in the two aforementioned metrics ($R_{\%}^{\text{max}}$ and $\langle R \rangle_{\%}$) which are tabulated below in Table 5.1. As the heavy element changes from Li to F in the first row, the position of the center-of-mass of the electron pair migrates substantially from the H to the heavy atom and this is captured very clearly within $E(R)$.

A second set of calculations were carried out on a set of truncated hydrides (i.e. the full molecular system was A−H). The remaining valence sites were left empty and the multiplicity was adjusted accordingly. This set of truncated hydrides was analyzed in order to determine what effect, if any, the presence of additional hydrogen atoms had on the extracule density of the A−H bond LMO, thereby differentiating between group and atomic electronegativities where the “group" consists solely of H substituents on the heavy atom of interest (Table 5.2).

Five separate electronegativity scales were chosen for comparison in this study. They are the Sanderson scale (I)[174], the Pauling Scale (II)[169], the Allred-Rochow scale (III)[175], the Allen Scale (IV)[180], and the Mulliken relation (V)[170]. These particular scales were chosen to cover a wide range of varying definitions for electronegativity. As previously noted, Pauling’s scale was based on the dissociation energies of homo and heteronuclear diatomics. Alternatively, Sanderson utilized a measure of compactness of an atom while the Allred-Rochow method related electronegativity to effective nuclear charge and covalent radius. The Allen Scale is based on configuration energies, which were defined as the average one-electron energy of a valence-shell electron in a single atom whereas Mulliken defined it simply as the arithmetic mean of the ionization potential and the electron affinity.

Table 5.1: Metrics of $E(R)$ for the A−H bond LMO in saturated hydrides.

			χ of atom A				
System	$R_{\%}^{\max}$	$\langle R \rangle_{\%}$	I	II	III	IV	V
LiH	0.977	0.978	0.886	0.980	0.970	0.912	0.970
BeH_2	0.895	0.873	1.810	1.570	1.470	1.576	1.540
BH_3	0.835	0.805	2.275	2.040	2.010	2.051	2.040
CH_4	0.745	0.732	2.746	2.550	2.500	2.544	2.480
NH_3	0.450	0.664	3.194	3.040	3.070	3.066	3.040
OH_2	0.378	0.607	3.654	3.440	3.500	3.610	3.680
FH	0.322	0.537	4.000	3.980	4.100	4.193	4.300
NaH	0.984	0.968	0.835	0.930	1.010	0.869	0.910
MgH_2	0.943	0.909	1.318	1.310	1.230	1.293	1.370
AlH_3	0.912	0.880	1.714	1.610	1.470	1.613	1.710
SiH_4	0.882	0.833	2.138	1.900	1.740	1.916	2.280
PH_3	0.842	0.772	2.515	2.190	2.060	2.253	2.410
SH_2	0.790	0.712	2.957	2.580	2.440	2.589	2.860
ClH	0.723	0.653	3.475	3.160	2.830	2.869	3.340
$R^2 : R_{\%}^{\max}$	- - -	- - -	0.775	0.853	0.914	0.888	0.810
$R^2 : \langle R \rangle_{\%}$	- - -	- - -	0.992	0.993	0.973	0.981	0.975

As previously noted, one might expect that the centre-of-mass of the electron pair within a bond LMO would provide an indication of electronegativity. Thus, using the two aforementioned metrics, correlations between each of the given scales and the two metrics were sought. The data for both the saturated and truncated hydrides are presented in Tables 5.1 and 5.2, respectively.

For both the saturated and truncated molecular systems, $R_{\%}^{\max}$ did not exhibit particularly compelling correlations with established electronegativity values, as is evident based on the coefficients of determination (R^2) being as low as 0.775. Visual evidence for this is provided in Figure 5.2 a) which depicts the relationship between Pauling electronegativities, χ_{Pauling} and $R_{\%}^{\max}$. This graph clearly demonstrates the presence of three outliers which represent the F−H, O−H, and N−H

Table 5.2: Metrics of $E(R)$ for the A−H bond LMO in truncated hydrides.

			χ of atom A				
System	$R_{\%}^{\max}$	$\langle R \rangle_{\%}$	I	II	III	IV	V
LiH	0.977	0.978	0.886	0.980	0.970	0.912	0.970
BeH	0.896	0.885	1.810	1.570	1.470	1.576	1.540
BH	0.826	0.795	2.275	2.040	2.010	2.051	2.040
CH	0.734	0.717	2.746	2.550	2.500	2.544	2.480
NH	0.434	0.656	3.194	3.040	3.070	3.066	3.040
OH	0.373	0.604	3.654	3.440	3.500	3.610	3.680
FH	0.336	0.552	4.000	3.980	4.100	4.193	4.300
NaH	0.984	0.968	0.835	0.930	1.010	0.869	0.910
MgH	0.935	0.898	1.318	1.310	1.230	1.293	1.370
AlH	0.924	0.890	1.714	1.610	1.470	1.613	1.710
SiH	0.878	0.828	2.138	1.900	1.740	1.916	2.280
PH	0.834	0.766	2.515	2.190	2.060	2.253	2.410
SH	0.786	0.708	2.957	2.580	2.440	2.589	2.860
ClH	0.723	0.653	3.475	3.160	2.830	2.869	3.340
$R^2 : R_{\%}^{\max}$	- - -	- - -	0.776	0.851	0.913	0.886	0.809
$R^2 : \langle R \rangle_{\%}$	- - -	- - -	0.989	0.987	0.968	0.975	0.969

bonds. These outliers are present for all 5 electronegativity scales and are the major factor for the poor relationships. Upon the removal of these outliers from the dataset, the R^2 value for the relationship with χ_{Pauling} improved from 0.853 to 0.965, along with substantial improvements in the fits for the other four scales.

As the average value of R is more indicative of the *full* distribution than simply the maximum, one might expect that the relationship between this property and electronegativities would be better. This is strongly supported by the data as when considering $\langle R \rangle_{\%}$, the correlation between it and established electronegativity metrics increased significantly and the F−H, O−H, and N−H bonds were no longer outliers. For this metric, the R^2 values ranged from 0.969-0.993, with the relationship for χ_{Pauling} depicted in Figure 5.2 b). As evidenced by this plot, there are no

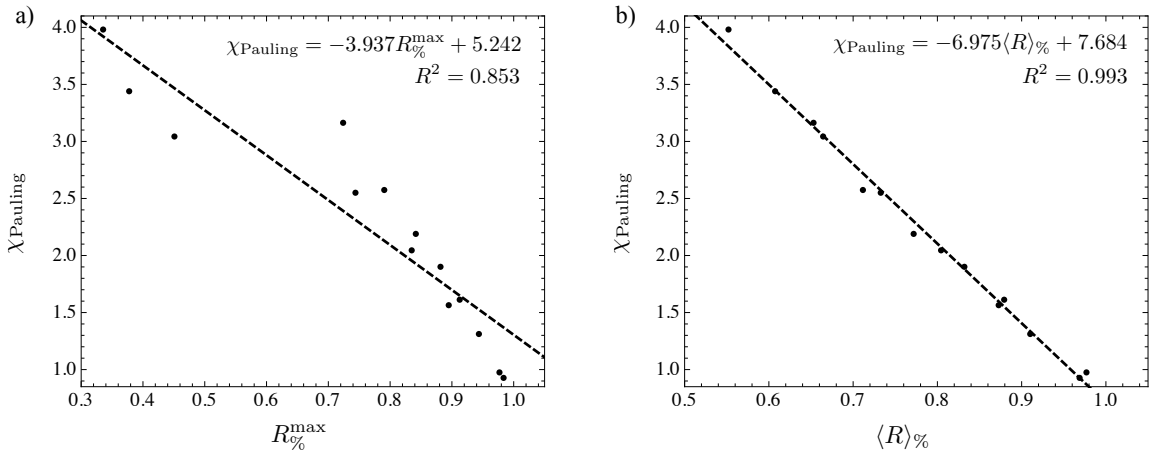


Figure 5.2: Correlation between Pauling electronegativities and a) $\langle R \rangle_{\%}$, and b) $R_{\%}^{\text{max}}$ for the first and second row hydrides.

points within the dataset that deviate significantly from the linear relationship.

What is more enlightening is to look at the differences between the metrics for the saturated and truncated hydrides. While systems with formulae A−H and A−H_2 are obviously very similar between these two sets, molecules with more valence sites begin to differ significantly in terms of our two metrics. This suggests that this method would be highly useful for defining group electronegativities, and based on the strong correlations with current electronegativity scales, the accuracy of such an approach would be quite high. One might be surprised that better correlations were obtained for the saturated molecular systems as opposed to the truncated hydrides despite the presence of additional hydrogen atoms that could distort the extracule density. However, as hydrogen atoms are considered as the reference for electron donating versus electron withdrawing, they would be expected to have minimal effects on the extracule density. They simply serve to saturate all valence sites on the heavy atom. Furthermore, the presence of radicals and or paired electrons that would otherwise be involved in bonding interactions is sure to introduce some error

into the modelling of the systems. Thus, we propose that saturating all valence sites of atom A with hydrogen atoms is a more accurate way to model the electronegativity for said atom. Conversely, if one were to substitute the hydrogen substituents with other functional groups, one could easily extend this method to determine electronegativities for the corresponding chemical group (e.g. -CH_3 vs -CH_2OH or -CH_2F).

In comparison to the topological method of Boyd and Edgecombe, which had a correlation coefficient of 0.991 (for χ_{Pauling}) using the power series relationship between χ_A and F_A, our approach works very well ($R^2 = 0.993$). Considering its simplicity, as one only need consider the average distance of the electron pair centre-of-mass from the atom of interest, this model is a simple, intuitive, and accurate approach to quantifying, perhaps even defining, electronegativity from a purely theoretical standpoint.

As the Pauling scale is so popular, many other electronegativity methods are modified to conform to the values of this scale. This is often done by scaling the electronegativity values of the new method to fit within the upper and lower bounds of the Pauling method (F as the upper bound and Cs or Fr as the lower bound). Applying a similar approach here, we use the linear relationship between χ_{Pauling} and $\langle R \rangle_{\%}$ (as shown in Figure 5.2 b) to develop our own electronegativity scale, χ_{LPM}. The electronegativity values obtained from this approach are compiled below in Table 5.3. The largest deviations in the χ_{LPM} values with respect to the Pauling values were for fluorine and sulfur at approximately 0.14 while the smallest deviation was observed for sodium (0.001). Regardless, the mean absolute deviation for the set of

Table 5.3: χ_{LPM} of the first and second row atoms

	χ_{LPM} by group						
System	1	2	13	14	15	16	17
First row	0.866	1.596	2.069	2.577	3.050	3.447	3.833
Second row	0.929	1.341	1.548	1.876	2.297	2.721	3.130

14 atoms was only 0.054. Furthermore, considering we have strong agreement between our model and the five electronegativity scales (that are significantly different in origin) studied herein, our model demonstrates strong potential for consideration as a novel, purely theoretical, method for describing electronegativity based on a novel topology of the electron pair density of single electron pairs.

As the truncated hydrides did not correlate as well as the saturated systems, only the saturated systems were analyzed with respect to the position intracules. As before, all first and second row hydrides were geometry optimized at the HF/u-6-311G(d,p) level of theory and the localized molecular orbitals were determined using the Edmiston-Ruedenberg model. $P(u)$ was then calculated for each of the saturated hydrides for the A−H bond LMO. The results of these calculations are tabulated below in Table 5.4

Table 5.4: R^2 values for metrics of $P(u)$ for the A−H bond LMO in saturated hydrides.

Metric	I	II	III	IV	V
u^{max}	0.949	0.960	0.965	0.972	0.931
$\langle u \rangle$	0.925	0.911	0.895	0.925	0.887
$\langle u^{-1} \rangle$	0.888	0.938	0.973	0.971	0.912
$u^{\text{max}}_{\%}$	0.005	0.003	0.000	0.003	0.024
$\langle u \rangle_{\%}$	0.001	0.000	0.001	0.000	0.006
$\langle u^{-1} \rangle_{\%}$	0.808	0.875	0.933	0.915	0.826

The data demonstrates that better correlations to existing electronegativity scales were observed for the extracule densities. As discussed, this is not overly surprising considering that the extracule describes more absolute position information while the intracules only contain relative position information. Nonetheless, strong correlations are still observed for each of the metrics studied for the intracule density.

Unlike for $E(R)$, the metrics for $P(u)$ perform far better when they are not scaled by the bond length. One possible reason for this is the relative nature of the position information. With the extracule density, observing how far the density shifts towards the electronegative atom will be significantly affected by the distance between the two bonded atoms; however, it should not be surprising that the relative positions of the electrons are not affected by this bond distance. Thus, scaling these metrics by this bond distance appears to remove any correlation between the metrics (especially the average and maximum positions of u) and the electronegativity scales.

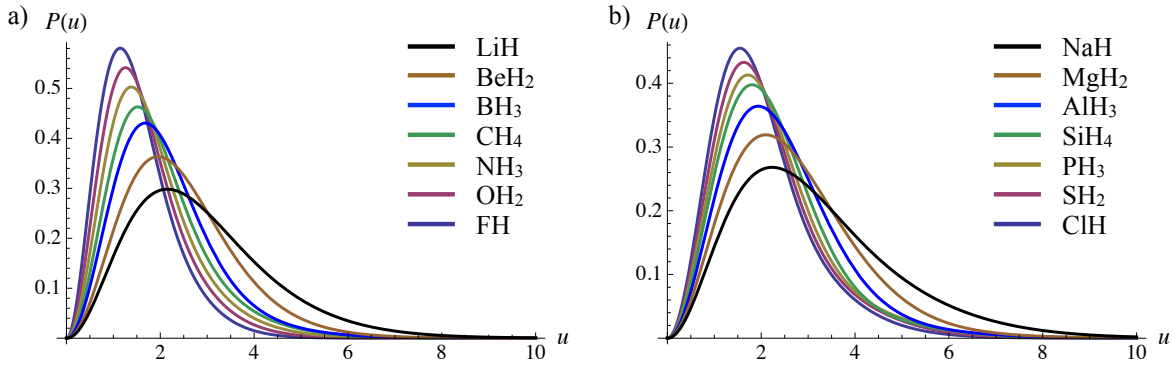


Figure 5.3: Localized intracule densities for the A−H bond in saturated a) first row hydrides, and b) second row hydrides.

5.3.2 Density Functional Theory

As discussed in the introductory chapter of this thesis, density functional theory (DFT) is a method that utilizes the one-electron density, $\rho(\mathbf{r})$ as opposed to the molecular wavefunction, $\Psi(\mathbf{r_1}, \ldots, \mathbf{r_N})$. Theoretically, this method is exact is the proper functional is known; however, this is clearly not the case and there have been numerous functionals developed over the years to try and obtain better results by obtaining the exchange and correlation energies to greater accuracy. In the remainder of this study on electronegativity, we decided to explore the effects of using density functional theory to observe whether greater correlations to electronegativity scales were obtained by using these *more accurate* methods.

For the purposes of this study, four separate DFT methods were chosen: BLYP, B3LYP, B3PW91, and M06-2X. While any of the countless DFT methods could have been studies, these four were specifically chosen due to the availability of electronic energies for each of the first and second row hydrides in the NIST database[184] and compatibility with the GAMESS software package which is used to calculate the LMOs.

To assess the accuracy of each of the four DFT methods, two separate metrics were used, both related to electronic energies. First, the mean absolute error (MAE) of each method was determined by comparing the energy of each hydride calculated by that method to the *exact* energy (defined as CCSD(T)/cc-pCVTZ). Second, as it is important to accurately describe the entire set under investigation (in this case, all of the first and second row hydrides), a second metric was employed that con-

Table 5.5: Accuracy metrics for the four DFT methods and HF.

Metric	HF	BLYP	B3LYP	B3PW91	M06-2X
MAE (E_h)	0.4183	0.0717	0.0519	0.0596	0.0556
MAE$_{max}$ (E_h)	1.1224	0.2117	0.1612	0.1645	0.1773
s (E_h)	0.2003	0.0587	0.0429	0.0430	0.0478
RSD	47.89%	81.81%	82.53%	72.12%	86.06%

sidered the largest absolute error encountered for each method. This metric, which we will call MAE$_{max}$, is defined as

$$\text{MAE}_{max} = \text{MAE} + \text{Max}(|E_i^{\text{exact}} - E_i^{\text{DFT}}) \tag{5.14}$$

where i is an index relating to each of the hydrides in the set.

The results of the accuracy tests on each of the DFT methods (as well as those from the HF method) are provided in Table 5.5. Also in this table are the standard deviation, s, and the relative standard deviation, RSD, which will be discussed later in this section. Based on both MAE and MAE$_{max}$, B3LYP is the most accurate functional on this test set. Based on MAE, the accuracy decreases in the order

B3LYP > M06-2X > B3PW91 > BLYP > HF

This order is slightly different when considering the alternative metric, MAE$_{max}$, where B3PW91 exhibits a lower max error resulting in a lower MAE$_{max}$. The rest of the set remains in the original order.

The extracules for all of the saturated hydrides were then calculated and the results are displayed in Figure 5.4. Visually, there are no noticeable differences

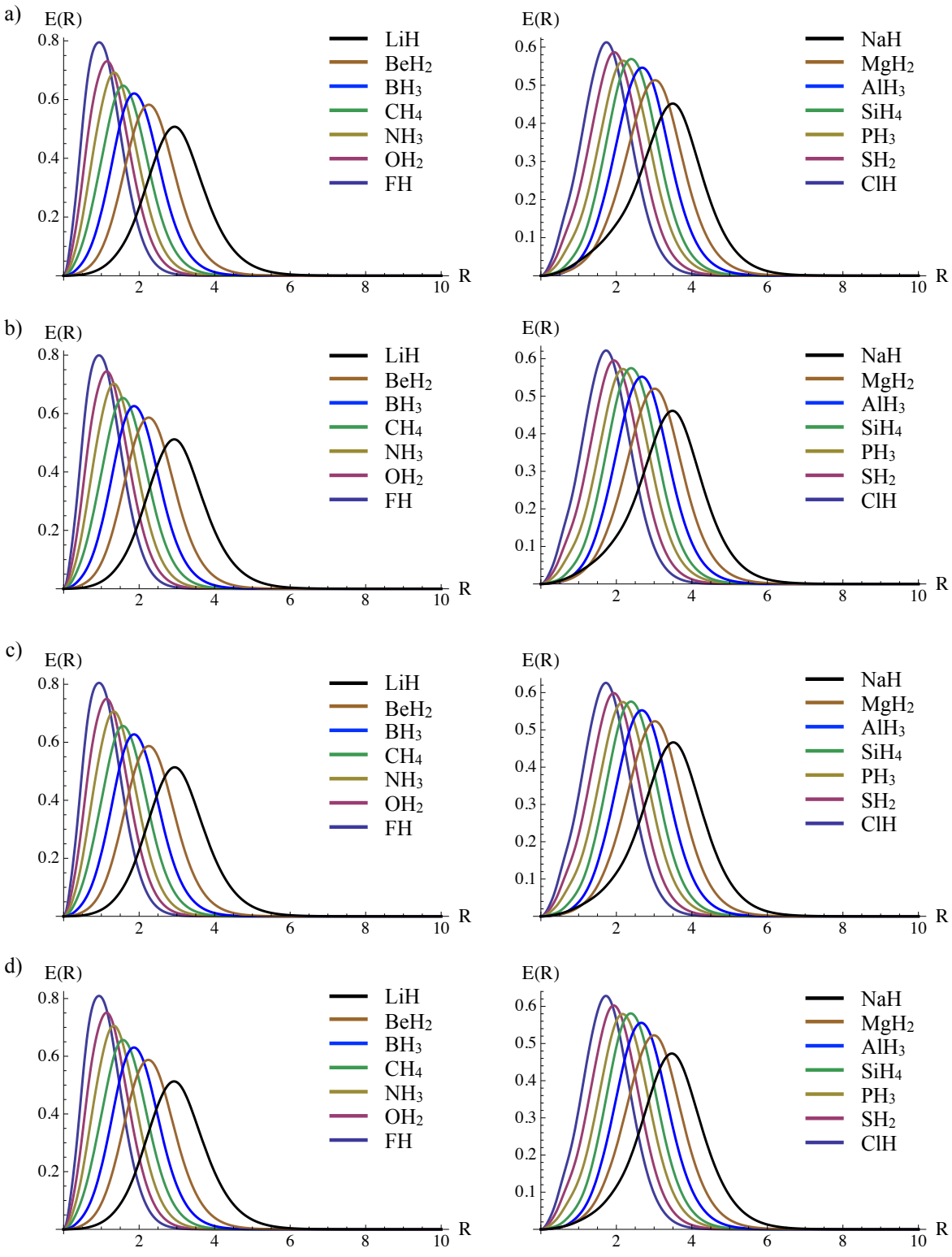


Figure 5.4: Localized extracule densities for the A−H bond in saturated first row (left) and second row (right) hydrides calculating using a) BLYP, b) B3LYP, c) B3PW91, and d) M06-2X.

Table 5.6: Coefficient of determination comparison for each computational method.

Metric	Scale	HF	BLYP	B3LYP	B3PW91	M06-2X
$R_{\%}^{\max}$	I	0.775	0.785	0.785	0.756	0.788
	II	0.853	0.861	0.862	0.833	0.864
	III	0.914	0.920	0.922	0.902	0.924
	IV	0.888	0.895	0.896	0.879	0.898
	V	0.810	0.822	0.823	0.795	0.823
$\langle R \rangle_{\%}$	I	0.992	0.984	0.986	0.973	0.991
	II	0.993	0.983	0.987	0.975	0.988
	III	0.973	0.958	0.963	0.966	0.963
	IV	0.981	0.967	0.972	0.982	0.975
	V	0.975	0.978	0.980	0.970	0.980

for the position extracules for each of the DFT methods. However, there are some minor quantitative differences. As before, $R_{\%}^{\max}$ and $\langle R \rangle_{\%}$ were calculated for each of the extracules. The coefficients of determination between these metrics and the five electronegativity scales are tabulated below in Table 5.6. In the table, the electronegativity scale (rows) and computational method (columns) that had the strongest overall correlation to each metric are highlighted in green.

Based on the methods that correlate best with existing electronegativity scales (M06-2X for $R_{\%}^{\max}$ and HF for $\langle R_{\%} \rangle$), there does not appear to be a connection between the accuracy of the model for predicting electronic energies and the relationship to electronegativity. One could point to the lower RSD for the HF method to suggest why it performs so well, but this would be a loose connection. Further investigation is required to determine why certain methods perform better than others. Nonetheless, all of these relationships are based off data from existing electronegativity scales, which, based on the number of different scales in existence, are flawed in their own right. Thus, even though these DFT methods did not agree

with existing scales better than did the HF method, this is not to say that they do not provide a more accurate theoretical description of the property of electronegativity.

5.4 Conclusions

We have applied a localized position intracule and extracule analysis to the set of first and second row hydrides in an attempt to develop a relationship between the topology of electron pair distributions and the concept of electronegativity. The localized extracule density was an obvious choice for application to electronegativities based on the absolute electronic position information contained within the distribution. We have demonstrated excellent correlations to several pre-existing electronegativity scales and thus, propose that the aforementioned approach could be used as a novel method for defining electronegativity in a simple, intuitive, and accurate way based on topological properties of the electron pair density. While the results obtained from $P(u)$ weren't quite as good as those from $E(R)$, strong correlations were still obtained. Analysis with various DFT functionals did not appear to increase the relation to existing electronegativity scales despite the increased accuracy of the models. This specific study illustrates another useful application of the SEPDA package in addition to those that we have highlighted in previous reports.[82,83,91]

Chapter 6

Using the Single Electron Pair Distribution Analyzer to Describe the Nature of the Hydrogen Bond

This chapter has been reproduced with modifications with permission from Proud, A.J.; Sheppard, B.J.H.; Pearson, J.K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 20194-20204. All work reported in this chapter was performed by Proud; however, some of the intracule calculations had been a repetition of work conducted by Sheppard. It was repeated to ensure consistency between the extracule and intracule data.

6.1 Introduction

'*. . . under certain conditions an atom of hydrogen is attracted by rather strong forces to two atoms, instead of only one, so that it may be considered to be acting as a bond between them. This is called the hydrogen bond.*' - Linus Pauling[185]

While our understanding of the hydrogen bond has improved significantly since it was defined by Pauling, his original definition does accurately describe the unique interaction. A hydrogen atom, while covalently bound to one electronegative atom, may interact with a second atom within a separate molecule or a separate moiety within the same molecule. The strength of this interaction depends on various factors including molecular geometries and electrostatics but can vary from 0.5-45 kcal/mol.[88]

Hydrogen bonding is highly prevalent in nature through its extensive involvement in the secondary structure of proteins, the stabilization of the double helix structure of DNA, and in many of the remarkable properties of H_2O. It was first described in the 1920s and its definition continues to change as our understanding of the interaction improves.[89] While it was once largely considered an electrostatic interaction, recently a resonance-covalency model has begun to emerge as a more complete characterization.[88,89]

As hydrogen bonding involves the through-space interaction of a lone pair of electrons of a heavy atom (Y) with an electron deficient neighbouring hydrogen (**H**−X), observing the changes in electron positions and electron pair separations represent intuitive ways to analyze and further characterize hydrogen bonding in-

teractions. While the wavefunction contains all necessary information regarding the positions of all N electrons for this purpose, dealing with such vast amounts of information is not generally tractable. However, as noted, hydrogen bonds primarily involve interactions between two localized sets of electron pairs: those involved in the covalent X−H bond, σ_{XH}, of the hydrogen bond donor and those in the lone pair of atom Y, n_Y, of the hydrogen bond acceptor. Thus, this interaction can be simplified by analyzing *individual* electron pairs. To do so, we can utilize the SEPDA package, notable the position intracule and extracule. For the purposes of this hydrogen-bonding analysis, it would be fruitful to study the 3-D (or 2-D) position extracule as opposed to the scalar counterpart. While this does significantly increase the computational cost of the calculations, it should afford a richer analysis of the hydrogen bonding interaction making this added time worthwhile.

By using LMOs, we can directly analyze the electron pair behaviour in molecular systems, and specifically the hydrogen bonding interaction (i.e. the σ_{XH} and n_Y orbitals) within the present work to observe how the centre-of-mass of the electrons described by these LMOs changes as the hydrogen bonding complex forms. Subsequently, we may also explore how such a density differs within different hydrogen bonding environments.

6.2 Computational Methods

Localization algorithms can be applied to any single determinant method, such as Hartree-Fock theory, which we explored in previous work describing the use of extracules in localized space,[91] or as will be explored here, Kohn-Sham (KS)

DFT. What remains to be seen is whether analyzing a property of electron pairs is feasible using a technique which only considers the one-electron density, $\rho(\mathbf{r})$. The traditional electron density is defined as

$$\rho(\mathbf{r}) = \int |\Psi(\mathbf{x}, \mathbf{x_2}, \ldots, \mathbf{x_N})|^2 \mathrm{d}s \mathrm{d}\mathbf{x_2} \ldots \mathrm{d}\mathbf{x_N} \tag{6.1}$$

A two-electron, single determinant wave function can be obtained from a single localized molecular orbital, ψ_k. Using this definition in place of the molecular wave function in equation (6.1), one obtains the spin-reduced two-electron density for the pair of electrons described by that specific LMO, $\rho^k(\mathbf{r_1}, \mathbf{r_2})$. For a single localized molecular orbital, no exchange interactions exist. In this scenario and defining the electron density for the specific LMO, k, as was done for $\rho^k(\mathbf{r_1}, \mathbf{r_2})$, the localized pair density is given by:

$$\rho^k(\mathbf{r_1}, \mathbf{r_2}) = \rho^k(\mathbf{r_1})\rho^k(\mathbf{r_2}) \tag{6.2}$$

In doing so, Coulombic electron correlation effects are being ignored. This omission is implicitly accounted for to a certain extent through the correlation functional, but it is still an approximation.

Recall that the extracule density is obtained from the pair density by the following relationship:

$$E(\boldsymbol{R}) = \int \rho(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta(\boldsymbol{R} - \tfrac{|\boldsymbol{r_1} + \boldsymbol{r_2}|}{2}) \mathrm{d}\boldsymbol{r_1}\, \mathrm{d}\boldsymbol{r_2} \tag{6.3}$$

Substituting equation (6.2) into (6.3) yields

$$E(\mathbf{R}) = \int \rho^k(\mathbf{r})\rho^k(2\mathbf{R} - \mathbf{r})\mathbf{dr} \tag{6.4}$$

If we define the molecular orbitals as a linear combination of Gaussian primitives, we obtain:

$$E(\mathbf{R}) = \sum_{\mu\nu\lambda\sigma}^{K} c_{\mu,k}c_{\nu,k}c_{\lambda,k}c_{\sigma,k}(\mu\nu\lambda\sigma)_E \tag{6.5}$$

where $c_{\mu,k}$ is the atomic orbital coefficient describing how the contribution of basis function ϕ_μ to the LMO of interest, ψ_k. These integrals can be determined from the recurrence relation developed by Thakkar and Moore[75] and the fundamental integral which was previously defined as

$$(\mu\nu\lambda\sigma)_E = \left(\frac{4\pi}{\zeta+\eta}\right)^{3/2} \exp\left[-\frac{\zeta\eta(2\mathbf{R}-\mathbf{P}-\mathbf{Q})^2}{\zeta+\eta}\right] \times \exp\left[-\frac{\alpha\beta(\mathbf{A}-\mathbf{B})^2}{\zeta} - \frac{\gamma\delta(\mathbf{C}-\mathbf{D})^2}{\eta}\right] T_x T_y T_z \tag{6.6}$$

The value of the extracule density was calculated at various points defined by the Mura-Knowles grid that was adapted to extend in the negative directions of the Cartesian coordinates. The grid was overlaid on the yz-plane and consisted of 151 points (75 in each direction as well as the origin). For the purposes of this study, the two atoms involved in the bond (X-H) containing the hydrogen atom as well as the atom containing the lone pair (Y) that will participate in hydrogen bonding are all positioned in the yz-plane to ensure adequate sampling of the environments around these atoms.

To quantify the extracules, there were a number of metrics employed. Since slices were chosen as opposed to all of 3-D space, the extracule did not integrate to unity (i.e. number of electron pairs; a necessary condition for single orbitals), and instead is equal to the value of $E(\boldsymbol{R_{yz}})$ integrated over the full slice. We define this metric as the zeroth moment, $\langle \boldsymbol{R^0_{yz}} \rangle$, which is defined as

$$\langle \boldsymbol{R^0_{yz}} \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(\boldsymbol{R_{yz}}) \mathrm{d}R_y \mathrm{d}R_z \tag{6.7}$$

Similarly, as was done previously, we can look at the first moment of R_z to determine the average value in the bond axis direction (z-axis). This is evaluated as

$$\langle R_z \rangle = \frac{1}{\langle \boldsymbol{R^0_{yz}} \rangle} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_z \times E(\boldsymbol{R_{yz}}) \mathrm{d}R_y \mathrm{d}R_z \tag{6.8}$$

The final quantifiable properties of interest in this study are the position of the extracule density maximum in the bond axis, R_z^{max}, and the value of the function at this maximum, $E(R_{yz}^{\text{max}})$.

Each of these properties can be calculated for a hydrogen bonding complex without needing the extracule density for either the X−H or Y molecule by itself. However, one can imagine that comparing these complexes to the molecules in the absence of the hydrogen bonding interaction would be rather insightful. To do this, we explore the extracule deformation density, $\Delta E_d^{\phi}(\mathbf{R})$, for orbital ϕ which is defined as

$$\Delta E_d^{\phi}(\mathbf{R}) = E_d^{\phi,\text{complex}}(\mathbf{R}) - E^{\phi,\text{molecule}}(\mathbf{R}) \tag{6.9}$$

The deformation density is useful in demonstrating the effect on the extracule density as the hydrogen bonding interaction forms and how it changes as the distance between the two molecules increases. The main property of interest involving the deformation density is δ_R which is the magnitude of the difference between the complex and single molecule. This is defined as

$$\delta_R = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\Delta_d^{\phi} E(\boldsymbol{R_{yz}})| \mathrm{d}R_y \mathrm{d}R_z \tag{6.10}$$

All geometries for the hydrogen bonding complexes were obtained from Hobza's X40x10 and S66x8 data sets.[151,186] Single-point energy and orbital localization calculations were performed on each of these geometries at the M06-2X/u6-311G(d,p) level of theory (u indicates that the basis set was completely uncontracted) using the GAMESS software package.[20] The Mathematica 8 software package was utilized to interpolate the data obtained from grid-point evaluation for visual representation and further analysis.[156] Atomic units are used throughout this chapter unless otherwise stated.

Additional calculations utilizing the "Atoms in Molecules" (AIM) technique were carried out on these systems to compare the capabilities of extracules in the localized pair model to AIM theory with regards to interpreting hydrogen bonding interactions. Past research[152,153,187–190] has shown the utility of AIM theory in the study of hydrogen bonding interactions and thus, it should represent a good case for comparing the results of our novel approach.

6.3 Results and Discussion

For the purpose of this study, 28 hydrogen bonding complexes were chosen from the X40x10 and S66x8 data sets.[151,186] For each of the systems the extracule, $E(\boldsymbol{R_{yz}})$, density was calculated for the σ_{XH} LMO of the proton donor individually and in the X$-$H$\cdots$Y hydrogen bonding complex at H$\cdots$Y separations of $0.90d_0$, $0.95d_0$, $1.00d_0$, $1.05d_0$, $1.10d_0$, $1.25d_0$, $1.50d_0$, and $2.00d_0$, where d_0 is the geometry optimized distance between the two atoms (see Figure 6.1). Likewise, for n_{Y}, $E(\boldsymbol{R_{yz}})$ was calculated for the complex at each of the separations listed above, as well as that for the isolated proton acceptor, Y.

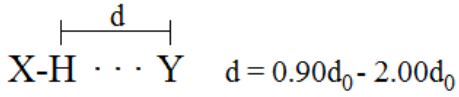


Figure 6.1: Pictorial representation of the various geometries for each of the hydrogen bonding complexes.

6.3.1 Extracule densities for σ_{XH} LMOs

When a hydrogen bond forms, the n_{Y} LMO interacts with the electron deficient H atom in the X$-$H molecule. In doing so, one expects that the electron pair, and thus the centre-of-mass, within the X$-$H bond would migrate towards the X atom due to the donation of election density from the lone pair to the hydrogen. This effect can be observed through both $\langle R_z \rangle$ and R_z^{max}. In the present molecules, the X$-$H bond midpoint is placed at the origin, with X in the negative z-direction and the hydrogen in the positive direction. The data presented in Table 6.1 clearly show that as Y approaches the X$-$H molecule, the average value of R_z shifts towards the

Table 6.1: $\langle \boldsymbol{R}_{yz} \rangle$ for the σ_{O-H} in MeOH-Y complexes

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
MeOH-CH_3Cl	0.0184	0.0223	0.0253	0.0275	0.0291	0.0319	0.0345	0.0359	0.0365
MeOH-CH_3F	0.0102	0.0151	0.0190	0.0220	0.0244	0.0287	0.0324	0.0349	0.0365
MeOH-$MeNH_2$	-0.0278	-0.0194	-0.0123	-0.0063	-0.0012	0.0098	0.0195	0.0274	0.0323
MeOH-MeOH	0.0043	0.0117	0.0178	0.0229	0.0271	0.0355	0.0424	0.0478	0.0511
MeOH-Peptide	-0.0091	-0.0008	0.0061	0.0120	0.0170	0.0275	0.0375	0.0461	0.0495
MeOH-Pyridine	-0.0273	-0.0184	-0.0109	-0.0027	0.0019	0.0121	0.0216	0.0283	0.0329
MeOH-H_2O	0.0044	0.0111	0.0166	0.0211	0.0247	0.0321	0.0384	0.0435	0.0469

X atom, as predicted. This effect is seen universally across the data set as shown in the table and in the Electronic Supplementary Information (ESI). The presence of the Y lone pair causes a small but significant migration of the centre-of-mass in the $\sigma_{\text{X-H}}$ bond towards the X atom. This effect is also observed in R_z^{max} which is included in the ESI.

In our previous work introducing the localized extracule, we noted a few trends in $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$.[91] First, the primary effect was that by introducing electronegative species to the neighbouring chemical environment, increases in $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ were observed in nearly all cases. This signifies that the value of the extracule density, $E(\mathbf{R})$, in the plane of the hydrogen bond increases as the electronegativity of neighbouring substituents increases indicating a migration of the centre-of-mass density towards the bonding region. The only exceptions were observed in systems where a second row element was bonded to a first row element. A similar trend is observed for these hydrogen bonding systems (Table 6.2). All systems exhibit an increase in $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ as the proton acceptor Y is drawn closer to the donor species, with the exception of systems where the donor contains an X$-$H bond with X=Cl,Br.

Within the context of AIM theory, an electron density analysis technique, values of the electron density at bond critical points ($\rho(r_c)$) have yielded direct relation-

Table 6.2: $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ for the $\sigma_{\text{X-H}}$ bond in a few select systems

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
HBr-MeOH	0.6620	0.6632	0.6640	0.6646	0.6651	0.6663	0.6674	0.6679	0.6682
HCl-$MeNH_2$	0.7176	0.7191	0.7205	0.7217	0.7227	0.7247	0.7262	0.7272	0.7271
HCl-MeOH	0.7340	0.7347	0.7351	0.7353	0.7356	0.7368	0.7368	0.7368	0.7368
HF-$MeNH_2$	1.0749	1.0712	1.0681	1.0655	1.0631	1.0576	1.0526	1.0498	1.0480
HF-MeOH	1.0724	1.0697	1.0673	1.0652	1.0634	1.0597	1.0572	1.0553	1.0541
$MeNH_2$-$MeNH_2$	0.8417	0.8408	0.8401	0.8395	0.8391	0.8385	0.8380	0.8374	0.8371
$MeNH_2$-MeOH	0.8413	0.8406	0.8401	0.8397	0.8394	0.8389	0.8384	0.8378	0.8375

ships to hydrogen bond strengths. This led us to explore whether any of the metrics employed herein would show similar predictive capacity. Figure 6.2 depicts the relationship between hydrogen bond strengths to $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ for each of the systems at equilibrium (i.e. $d = 1.00d_0$). While there does not appear to be any mathematical relationship, the graph clearly classifies the systems into 5 separate classes where the interaction energy may vary, but the zeroth moment is effectively unchanged.

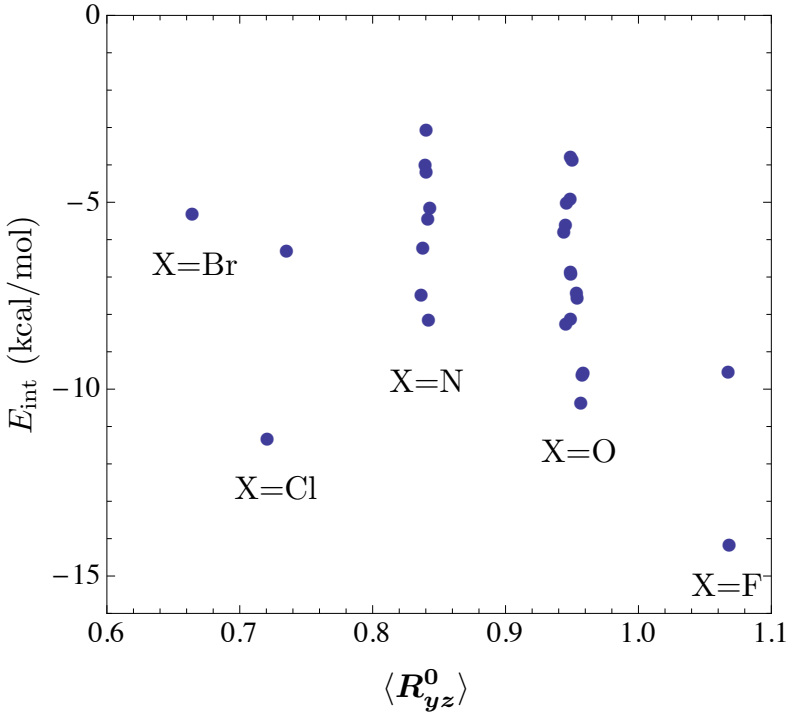


Figure 6.2: Classification of hydrogen bonding systems based on $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ of the $\sigma_{\text{X-H}}$ LMO and hydrogen bonding strengths.

Each class corresponds to a particular X atom in the X−H bond. In increasing order of $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$, these classes represent Br−H, Cl−H, N−H, O−H and F−H bonds. This ordering agrees with the previously noted relationship between $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ and electronegativity. The lone system containing Br (the lowest electronegativity) falls at the lower end of $\langle \boldsymbol{R}^{\mathbf{0}}_{\boldsymbol{yz}} \rangle$ values, while increases in electronegativity lead to an increase in the value of the metric.

The extracule deformation density is a useful visual tool as it highlights all of the trends that have been discussed thus far. Figure 6.3 illustrates $\Delta E^{\phi}_{1.00d_0}(\boldsymbol{R}_{\boldsymbol{yz}})$ for the $\sigma_{\text{N-H}}$ bond and n_{Y} LMOs in the $MeNH_2$-MeOH hydrogen bonding complex at its equilibrium geometry. This figure clearly demonstrates that the centre-of-mass in the $\sigma_{\text{N-H}}$ bond migrates away from the hydrogen atom towards the X atom. This migration is caused by the shift in the centre-of-mass from the lone pair, n_{O} towards

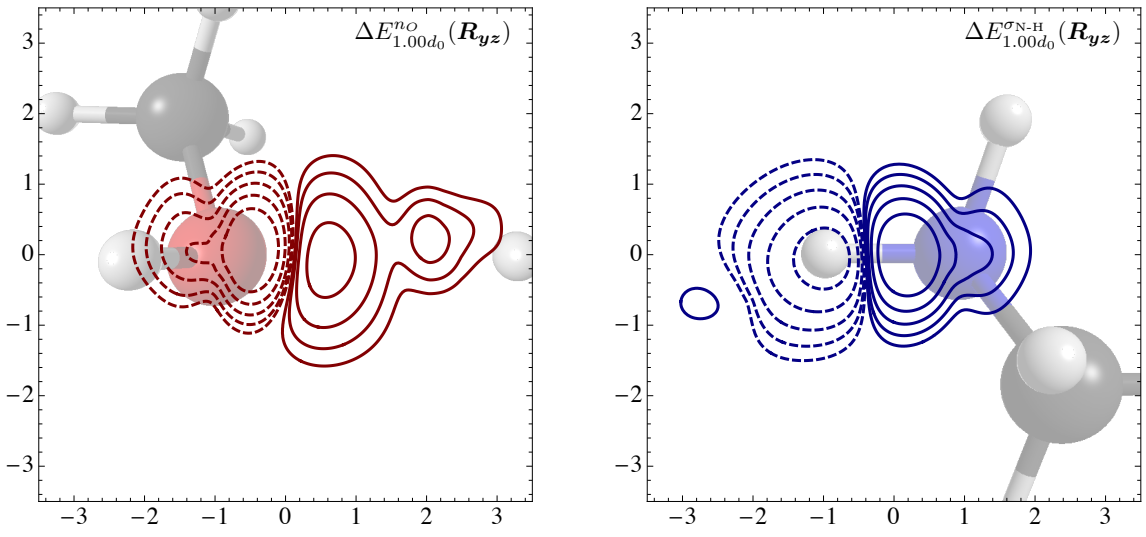


Figure 6.3: Contour plot of $\Delta E^{\sigma_{\text{N-H}}}_{1.00d_0}(\boldsymbol{R}_{\boldsymbol{yz}})$ and $\Delta E^{n_{\text{O}}}_{1.00d_0}(\boldsymbol{R}_{\boldsymbol{yz}})$ in $MeNH_2$-MeOH depicting the change in the extracule density in the presence of the proton acceptor, MeOH (dashed lines signify negative contours while solid lines signify the positive contours).

the hydrogen in X-H during the formation of the hydrogen bond. This agrees with the previous observation that both $\langle R_z \rangle$ and $R_z^{\max}$ shift towards X when the hydrogen bond forms and as the hydrogen bonding distance decreases.

In comparing Figures 6.4(a) and (b), it is noted that as the two molecules involved in the hydrogen bonding interaction separate, the extracule of the complex begins to resemble that of the single molecule in the absence of hydrogen bonding. The degree to which the extracules differ can be measured using δ_R. This property is tabulated in Table 6.3 for the complexes where water is the donor species (data for all systems can be found in the ESI). These data clearly show that a greater difference in the extracule density is observed when the two species are in close proximity. This difference tails off significantly as the systems reach a separation of $2.00d_0$. This again suggests a propensity for the localized extracule to offer, at minimum, qualitative insight into the strengths of the non-covalent interactions.

For a given complex, analysis of each quantitative metric described in the computational section demonstrates a strong quadratic relationship with respect to interaction energy (Figure 6.5). While it is expected that this relationship would not hold for larger separations, encountering such interactions in real chemical systems would be rare. Over the 28 systems analyzed, coefficients of determination

Table 6.3: δ_R for $\phi = \sigma_{\text{O-H}}$ in the $H_2O{-}Y$ subset of systems.

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$
H_2O-$MeNH_2$	0.0866	0.0735	0.0626	0.0535	0.0459	0.0298	0.0162	0.0060
H_2O-MeOH	0.0677	0.0564	0.0472	0.0397	0.0336	0.0213	0.0114	0.0043
H_2O-Peptide	0.1205	0.1109	0.1036	0.0981	0.0938	0.0860	0.0811	0.0795
H_2O-Pyridine	0.1156	0.1057	0.0983	0.0927	0.0885	0.0812	0.0772	0.0762
H_2O-H_2O	0.0149	0.0137	0.0108	0.0093	0.0081	0.0055	0.0031	0.0011

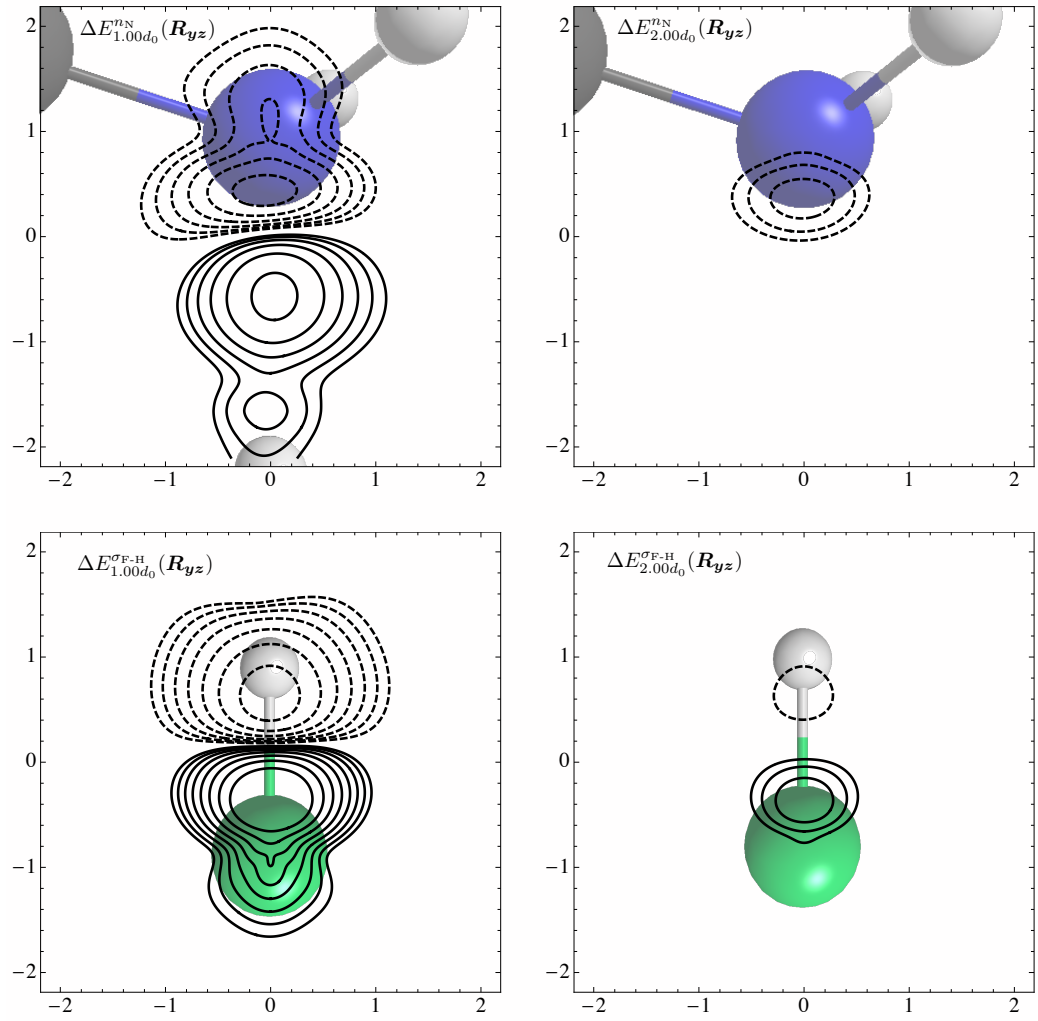


Figure 6.4: Contour plot of $\Delta E_d^{\sigma_{\text{F-H}}}(\boldsymbol{R_{yz}})$ and $\Delta E_d^{n_N}(\boldsymbol{R_{yz}})$for the HF-$MeNH_2$ complex demonstrating the diminishing effect on $\Delta E(\mathbf{R})$ as the complex separates (dashed lines signify negative contours while solid lines signify the positive contours).

(R^2) values ranged from 0.9113-0.9999. However, with the exceptions of H_2O-Pyridine (R^2=0.9113), H_2O-Peptide (R^2=0.9257), and Peptide-H_2O (R^2=0.9717) complexes, all other systems had R^2 values greater than or equal to 0.9941. In the case of the first two complexes, there may be some other form of interaction that is occurring in these systems that is causing these deviations from the quadratic relationship while the fit of the third appears to be distorted by a single data point.

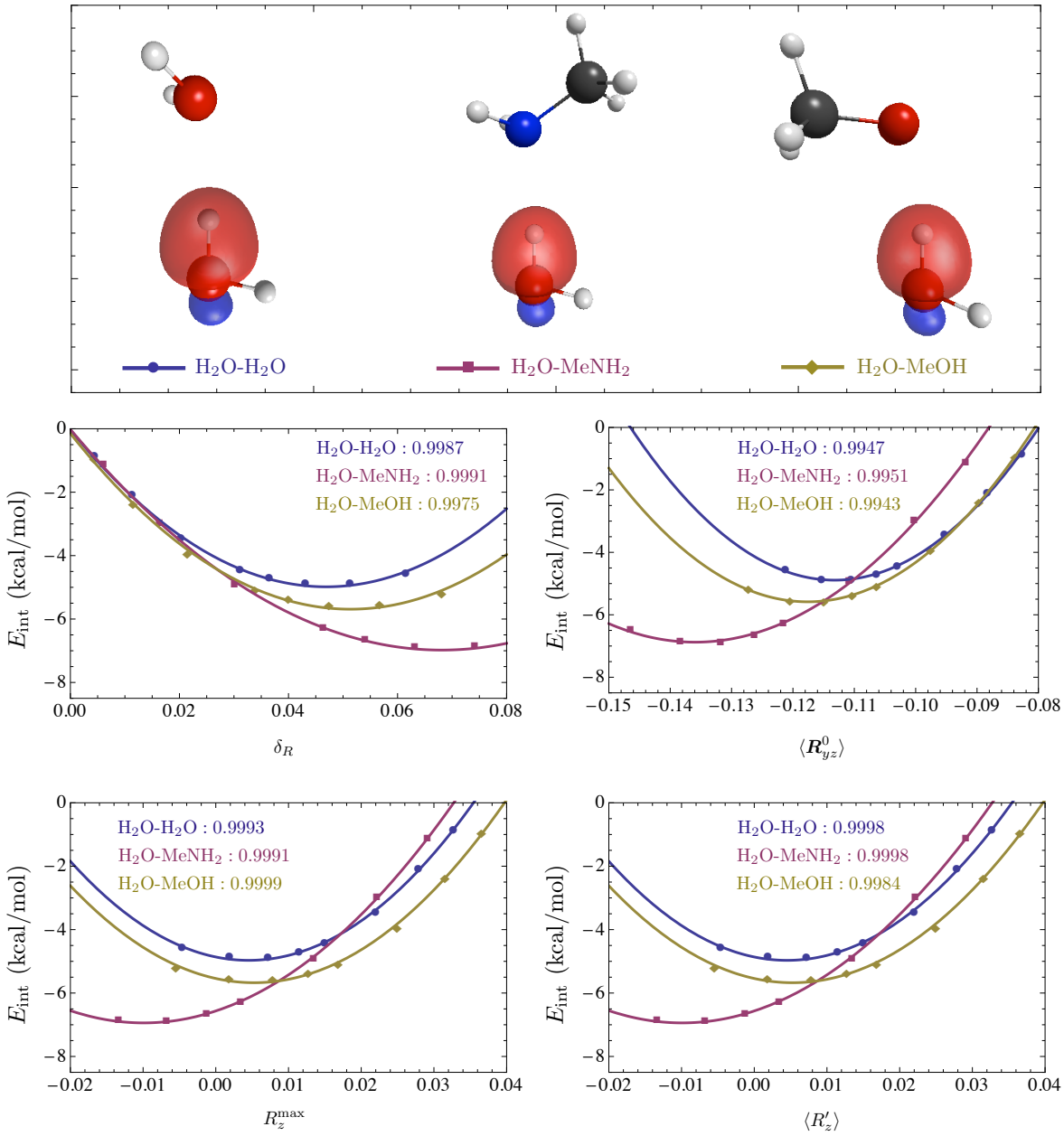


Figure 6.5: Relationship between δ_R, $\langle \boldsymbol{R}^0_{\boldsymbol{yz}} \rangle$, $R_z^{\max}$, and $\langle R'_z \rangle$ versus hydrogen bond strength for H_2O-H_2O, H_2O-$MeNH_2$, and H_2O-MeOH.

As in the case of $\langle \boldsymbol{R}^0_{\boldsymbol{yz}} \rangle$, we also explored the possibility of a relationship existing between δ_R for each of the systems when $d = 1.00 d_0$. This relationship is depicted in Figure 6.6. The two points in red represent the H_2O-Pyridine and H_2O-Peptide complexes. When these two points are removed, the coefficient of determination

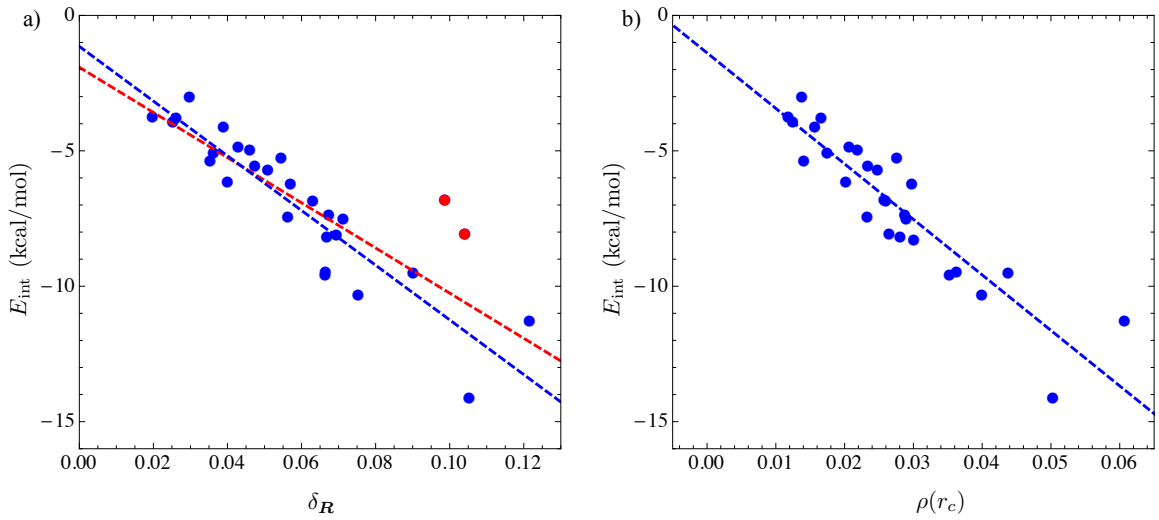


Figure 6.6: Relationship between hydrogen bond interaction strength, E_{int} versus a) δ_R, and b) the density at the bond critical point, $\rho(r_c)$, for all systems at equilibrium.

improved from 0.7137 to 0.8411. While this may not appear to have much predictive power, it is on par with bond-critical-point (BCP) density analysis of AIM theory (R^2=0.8282). Figure 6.6 b) demonstrates this analysis. Unlike the case of the extracules, there are no distinct outliers in this analysis that could lead to an improvement in the R^2 value. Past studies have shown that BCP densities are strong predictors for hydrogen bonding[152,153,187–190] suggesting that the variability in this data set makes predictions rather difficult. The fact that our technique performs as well as it does for this data set is a strong indicator of its potential in the analysis of hydrogen bonded complexes and non-covalent interactions in general.

6.3.2 Extracule densities for n_{Y} LMOs

As previously noted, there are two pairs of electrons involved in the hydrogen bonding interaction: the $\sigma_{\text{X-H}}$ bonding pair and the n_{Y} lone pair. The same analyses that were carried out in the previous section were also performed on the lone pair LMOs.

While analyses were carried out for both the geometry optimized structures and the varying separation distances between the hydrogen bond donor and acceptor, only the former will be discussed here as nothing of significance was obtained from the latter that hasn't been previously discussed.

The results of this analysis are tabulated below in Table 6.4. As before, the two outliers, H_2O-Peptide and H_2O-Pyridine were omitted from the analysis. Much like for the $\sigma_{\text{X-H}}$ bond LMO, the only metric that had any reasonable correlation to hydrogen bond interaction strengths, E_{int}, was δ_R. For the lone pairs, the coefficient of determination decreased to 0.5705 down from 0.8411 from the X-H bond LMOs, representing a significant decrease in predictive potential.

Unfortunately, these localized lone pairs proved to be far less predictive and thus we conclude that it is the $\sigma_{\text{X-H}}$ bond that is most important when characterizing the hydrogen bonding interactions. All other data, including those for the various separation distances, are tabulated in Appendix C with other relevant information from this chapter.

Table 6.4: Summary of the various metrics for each hydrogen bonding complex at equilibrium and the coefficient of determination for the relationship to E_{int}.

System	δ_R	$\langle \boldsymbol{R}^0_{yz} \rangle$	$R_z^{\max}$	$E(\boldsymbol{R}_{yz}^{\max})$	$\langle \boldsymbol{R}_{yz} \rangle$
$AcNH_2$-$AcNH_2$	0.0534	0.8742	0.4259	0.8134	0.4152
AcOH-AcOH	0.0806	0.8803	0.4263	0.8056	0.3958
CCl_3OH-H_2O	0.0700	0.8775	0.4255	0.8204	0.3814
CF_3OH-H_2O	0.0577	0.8783	0.4484	0.8288	0.4158
HBr-MeOH	0.0116	0.3683	0.8641	0.2706	0.7525
HCl-$MeNH_2$	0.1073	0.7585	0.3142	0.4800	0.2210
HCl-MeOH	0.0122	0.3692	0.8641	0.2712	0.7508
HF-$MeNH_2$	0.0790	0.7543	0.3165	0.4904	0.2748
HF-MeOH	0.0147	0.3714	0.9695	0.2688	0.7501
$MeNH_2$-$MeNH_2$	0.0264	0.7144	0.3891	0.4779	0.3957
$MeNH_2$-MeOH	0.0250	0.8782	0.4296	0.8611	0.4128
$MeNH_2$-Peptide	0.0185	0.7149	0.5741	0.6585	0.5943
$MeNH_2$-Pyridine	0.0205	0.7203	0.6288	0.5370	0.5808
MeOH-CH_3Cl	0.0138	0.6300	0.0604	0.2917	0.2101
MeOH-CH_3F	0.0314	1.0504	0.5027	1.3663	0.5051
MeOH-$MeNH_2$	0.0478	0.7466	0.3270	0.4953	0.3126
MeOH-MeOH	0.0365	0.8466	0.4297	0.7972	0.4327
MeOH-Peptide	0.0384	0.8091	0.4279	0.7496	0.4213
MeOH-Pyridine	0.0439	0.7299	0.3391	0.5320	0.2347
MeOH-H_2O	0.0347	0.8580	0.4411	0.8132	0.4377
Peptide-$MeNH_2$	0.0420	0.7394	0.3333	0.4958	0.3144
Peptide-MeOH	0.0067	0.3193	0.8802	0.2611	0.8330
Peptide-H_2O	0.0208	0.8760	0.6186	0.8517	0.6165
H_2O-$MeNH_2$	0.0442	0.7462	0.3365	0.4964	0.3249
H_2O-MeOH	0.0375	0.8680	0.4276	0.8317	0.4133
H_2O-H_2O	0.0354	0.8768	0.4346	0.8445	0.4183
R^2	**0.5705**	**0.0002**	**0.0104**	**0.0251**	**0.0736**

6.3.3 Subsets of Data

The results shown in Figure 6.2 suggest that it might be of interest to examine the data set based on the atom X for the $\sigma_{\text{X-H}}$ bond LMO and atom Y for the n_{Y} LMOs. Considering the full data set (minus the outliers), the only atoms, X and Y, that represent a reasonable sample size are the nitrogen and oxygen atoms. For the bond LMOs, there are a total of 8 complexes where X=N and 13 where X=O.

Similarly, for the lone pair in the acceptor species, there exists 9 systems where Y=N and 15 where Y=O. Thus, the relationship between the 5 metrics for each of these subsets to E_{int} were determined and are summarized in Table 6.5.

In this table, numbers shown in green indicate an improved relationship to E_{int} while those in red indicate a weaker relationship. For the $\sigma_{\text{X-H}}$ bond LMOs, considering the subsets led to significantly stronger correlations in nearly all cases. In fact, in many cases, especially for R_z^{max}, for the full data set, there is no relationship to hydrogen bond strengths, but very strong correlations when considering the identity of the atom X. The one exception is δ_R. The correlation to E_{int} for this metric is quite strong for the full set but only decreases marginally for the subsets.

As for the n_{Y} LMOs, aside from δ_R no strong relationships were observed for any metrics and any increases or decreases in the R^2 value are not of significance. However, once again, R^2 for δ_R decreased for Y=O. It is probable that this is caused by the positioning of the lone pairs of oxygen. In many cases, the lone pair is not oriented exactly at the hydrogen bond donor as it often is for Y=N, and thus, considering a slice of $E(\boldsymbol{R})$ could lead to deficiencies for these systems.

Table 6.5: Coefficients of determination for the relationship between each of the metrics and E_{int} for subsets of the full data set.

$\sigma_{\text{X-H}}$	δ_R	$\langle \boldsymbol{R}_{yz}^0 \rangle$	R_z^{max}	$E(\boldsymbol{R}_{yz}^{\text{max}})$	$\langle \boldsymbol{R}_{yz} \rangle$
X=All	0.8411	0.1022	0.2181	0.1633	0.3471
X=N	0.8172	0.0283	0.8416	0.4280	0.6263
X=O	0.5835	0.4495	0.8427	0.6499	0.7970
n_{Y}					
Y=All	0.5705	0.0002	0.0104	0.0251	0.0736
Y=N	0.7631	0.0853	0.3477	0.0020	0.4330
Y=O	0.3590	0.0003	0.0003	0.0041	0.0130

6.3.4 Scaling Metrics

Of the five metrics discussed herein, the one that consistently exhibited the strongest correlation to interaction energies was δ_R. Considering this metric is the only one that explicitly incorporates the change in the electron pair upon hydrogen bond formation, it was decided that a similar approach would be applied to the other 4 metrics. Thus, the change in each metric was observed with the calculation of the change determined as follows:

$$\Delta x = x^{\text{X-H}\cdots\text{Y}} - x^{\text{X-H}} \qquad \text{or} \qquad \Delta x = x^{\text{X-H}\cdots\text{Y}} - x^{\text{Y}} \tag{6.11}$$

Additionally, as a clear relationship was observed between the separation between the hydrogen bond donor/acceptor and the strength of the interaction (see Figure 6.5), considering the distance between the two moieties could be beneficial. For this reason, each of the metrics was also scaled by the bond length. This can be denoted as

$$x_{\%} = \frac{x}{d_0} \tag{6.12}$$

Combining both of these approaches would produce another scaling metric which will be denoted by $\Delta x_{\%}$. Each of these three scaling techniques were applied to the 5 metrics previously discussed. The results are summarized in Table 6.6. As before, improvements in the relationship, as determined by the coefficient of determination, are indicated by green text whereas deterioration is indicated by red. The standard for these comparisons are the unscaled metrics shown in Table 6.5.

Table 6.6: Coefficients of determination for the relationship between each of the scaled metrics and E_{int}.

Δx	δ_R	$\langle \boldsymbol{R}^0_{yz} \rangle$	$R^{\max}_z$	$E(\boldsymbol{R}^{\max}_{yz})$	$\langle \boldsymbol{R}_{yz} \rangle$
X=All	- - -	0.2625	0.1525	0.5173	0.2965
X=N	- - -	0.3909	0.7785	0.7284	0.7758
X=O	- - -	0.7157	0.4907	0.8169	0.8492
$x_{\%}$					
X=All	0.8954	0.5495	0.2262	0.3651	0.3231
X=N	0.8233	0.7384	0.8279	0.7400	0.4512
X=O	0.8782	0.7438	0.8637	0.7660	0.7881
$\Delta x_{\%}$					
X=All	- - -	0.3045	0.1700	0.5580	0.3501
X=N	- - -	0.4214	0.7584	0.7333	0.7869
X=O	- - -	0.7875	0.6606	0.8438	0.8901

Nearly all relationships improved upon scaling with the exception of R_{yz}. The reason for this is unclear and requires further investigation. Regardless, small, but meaningful improvements were observed for many of the other metrics under these scaling schemes. While δ_R remained one of the strongest metrics, especially for the full data set, strong correlations were observed for many of the scaled metrics.

The strongest correlation was observed by scaling the δ_R values by the hydrogen bond distance, the R^2 value improved to 0.8954 from 0.8411 and yielded the following relationship

$$E_{int} = -111.3\frac{\delta_R}{d_0} - 0.714 \tag{6.13}$$

6.3.5 Bivariate Analysis

Although analysis of the lone pair of species Y did not yield substantial quantitative information on its own, it was decided that analyzing the two localized electron pairs simultaneously might yield some added insight. Thus, bivariate analyses were

carried out with each of the quantitative metrics. While the regression results did improve, the improvements were marginal in most cases and not worth the added computational cost.

Thus, an alternative approach was explored. As discussed, a strong relationship was observed between δ_R and the hydrogen bond strength for each of the systems at equilibrium. Therefore, bivariate analyses were carried out with δ_R and each of the other metrics employed in this study. While the results did not improve dramatically, the R^2 values did approach, and in some cases surpass, 0.9.

6.3.6 Intracule Analysis

While the focus of this study was on the 3-D position extracule, it would not be unexpected for the position intracule to have potential for predicting hydrogen bond strengths. Under the resonance-covalency model[88,89] that was supported by the analysis of $E(\boldsymbol{R})$, one would expect that the electrons in the $\sigma_{\text{X-H}}$ bond LMO would contract (i.e. u would decrease) as the electron pair migrates towards X. Similarly, as the interaction between atoms Y and H form, the electrons in n_{Y} should separate leading to an increase in u. This is confirmed by the results shown for the H_2O dimer shown in Figure 6.7.

While all analyses carried out for $E(\boldsymbol{R})$ were also performed for $P(u)$, the results for the position intracule will be briefly summarized to prevent excessive repetition. All data pertaining to the analysis of $P(u)$ for these hydrogen bonding complexes is available in Appendix C.

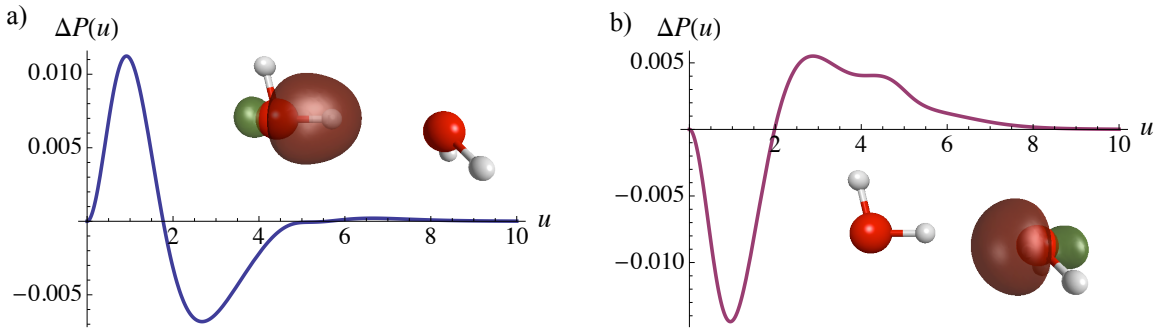


Figure 6.7: The deformation density of the position intracule for a) the $\sigma_{\text{O-H}}$ bond and b) the n_O lone pair, LMOs for the water dimer.

The results of the relationship between the metrics for $P(u)$ and E_{int} are shown in Tables 6.7 and 6.8. In these cases, scaling the metrics helps in nearly all cases. This may be due to the complete nature of $P(u)$ compared to the slices of $E(\boldsymbol{R})$ that were studied.

Much like in the case of the position intracule, the only unscaled metric that has any reasonable relationship to interaction energies is δ_u. Furthermore, this relationship is only evident for the $\sigma_{\text{X-H}}$ bond LMO and not the lone pair.

When considering the scaled metrics and the subsets of the full set; however, the position intracule does appear to be useful in predicting hydrogen bond strengths. Unlike the analyses conducted on $E(\boldsymbol{R})$, the scaled metrics make significant improvements to the correlations on the full data set.

While the strongest relationships to E_{int} were observed for $E(\boldsymbol{R})$, similar information can be drawn from the position intracule. All of the metrics when scaled demonstrate significant potential for predicting interaction energies, even with the lone pair LMOs. The difficulty with lone pairs in the extracular analysis, again, lies in the requirement to analyze slices as opposed to the full distribution.

Table 6.7: Coefficients of determination for the relationship between each of the scaled metrics and E_{int} for $\sigma_{\text{X-H}}$ bond LMO.

$\sigma_{\text{X-H}}$ / x	δ_u	$\langle u^{-1} \rangle$	u_{max}	$\langle u \rangle$
X=All	0.6647	0.0940	0.0573	0.0497
X=N	0.8345	0.5161	0.4760	0.4926
X=O	0.7499	0.6177	0.6647	0.2130
Δx				
X=All	- - -	0.5520	0.7580	0.4185
X=N	- - -	0.7931	0.7655	0.6740
X=O	- - -	0.7487	0.8180	0.6060
$x_{\%}$				
X=All	0.7399	0.5749	0.1720	0.1829
X=N	0.8336	0.7536	0.7494	0.7378
X=O	0.8152	0.7342	0.7098	0.7271
$\Delta x_{\%}$				
X=All	- - -	0.6186	0.8028	0.5092
X=N	- - -	0.8021	0.7786	0.7167
X=O	- - -	0.8031	0.8343	0.6871

Table 6.8: Coefficients of determination for the relationship between each of the unscaled/scaled metrics for $P(u)$ and E_{int} for n_{Y} LMO.

n_{Y} / x	δ_u	$\langle u^{-1} \rangle$	u_{max}	$\langle u \rangle$
X=All	0.2716	0.0417	0.0229	0.0358
X=N	0.2832	0.0223	0.0245	0.0521
X=O	0.2160	0.3626	0.5745	0.2159
Δx				
X=All	- - -	0.2948	0.7446	0.2201
X=N	- - -	0.4049	0.8645	0.1923
X=O	- - -	0.2696	0.6283	0.1057
$x_{\%}$				
X=All	0.3891	0.1738	0.6878	0.7476
X=N	0.4177	0.5800	0.8167	0.8177
X=O	0.3403	0.7306	0.7960	0.8084
$\Delta x_{\%}$				
X=All	- - -	0.4007	0.7838	0.3084
X=N	- - -	0.5107	0.8663	0.2700
X=O	- - -	0.3883	0.6811	0.2117

6.4 Conclusions

A novel method for the analysis of non-covalent interactions involving electron pairs, more specifically hydrogen bonding, was explored. This localized electron pair approach affords a quantitative analysis technique in accordance with the resonance-covalency model as a characterization tool for hydrogen bonding and general non-covalent interactions. A clear migration of the centre-of-mass of localized electron pairs was consistently observed towards the hydrogen with respect to the n_{Y} lone pair and away from the hydrogen in the case of the $\sigma_{\text{X-H}}$ bond. This, combined with the concomitant contraction of the $\sigma_{\text{X-H}}$ bond pair and expansion of the n_{Y} lone pair electrons strongly support the resonance-covalency model. Strong quadratic relationships were observed between each of the studied metrics and interaction energies for a given hydrogen bonding complex indicating a unique predictive quality of our characterization.

When considering the entire set of molecules, a thorough study was conducted on each of the systems at their equilibrium geometries. Of the metrics employed, δ_R and δ_u demonstrated the strongest relationship with respect to hydrogen bond strengths on the full set. This predictive power could be improved by studying the entire extracule density, as opposed to the single slices in the bond plane that were utilized here. The only issue with using δ_R as a predictive tool is that the extracule for the X$-$H molecule in the absence of the proton acceptor, Y, is required. However, by splitting the data set into the identity of the atoms X and Y or by scaling by bond length, strong relationships can be obtained that could be utilized for such cases.

Chapter 7

Conclusions & Future Work

7.1 Conclusions

The analysis of the electron pair properties of molecules has been focused on complete molecular properties for far too long. Herein, we have introduced a novel software package to study the properties of individual electron pairs. For most purposes, this includes localized pairs of electrons (i.e. those described by localized molecular orbitals and natural bond orbitals), but can be used in the study of an electron pair in a traditional delocalized molecular orbital (i.e. canonical molecular orbitals).

This new form of analysis affords a more detailed description of electron pair behaviour caused by specific chemical reactions. While earlier research in the Pearson lab demonstrated the ability to determine bond dissociation energies and relative acidities of benzoic acid derivatives from the position intracule, the applications of these analyses were greatly expanded in the present work. This software package is

capable of calculating intracules in position, momentum and combined (posmom) space in addition to a small set of extracules including both scalar and vectorized analogues in position space in addition to the momentum space distribution.

Much of the work presented herein focussed on the understudied 3-Dimensional position extracule in numerous different applications. The obvious benefit of examining the extracule in place of the intracule is the more absolute information regarding the electron pair. In position space, this provides a better idea of where in the system the electrons reside.

From both the vectorized and scalar forms of the position extracule, relationships to numerous chemical properties have been examined. We propose the use of the position extracule density as a novel, purely theoretical way to characterize electronegativity. In contrast to previous work using the electron density, a better, and much simpler relationship to existing electronegativity scales was obtained.

An extensive study on the effects of hydrogen bonding was also conducted with respect to both $E(\boldsymbol{R})$ and $P(u)$. While previous reports have demonstrated the power of AIM theory in describing any type of hydrogen bonding interaction, the data set employed in this study showed that this is not true in general. SEPDA was used to characterize hydrogen bonds and its capabilities were compared to those of AIM theory and it was demonstrated that the two techniques were on par. Improvements can be made to the single electron pair analysis by considering the atoms involved in the interactions as well as various different scaling parameters.

A number of unique bonding environments were analysed with the suite of tools available in the SEPDA package yielding interesting insight. These include the 3

centre-2 electron bonding environments seen in diborane amongst others as well as strained bonding environments observed in smaller cycloalkanes. The potential applications for SEPDA are only limited by the available chemistry. While we chose to represent commonly seen chemical interactions as well as a few interesting cases, this by no means represents a comprehensive study of the information that can be obtained by this software.

7.2 Future Work

There remains much to do in the area of studying localized electron-pair behaviour. As demonstrated throughout this thesis, SEPDA is a great tool for tackling this problem; however, the applications that have been explored thus far have only scratched the surface. Upon the completion of this thesis, a collaboration with Deslongchamps and Deslongchamps was being carried out to apply the SEPDA package to their research interests.

Herein, we mainly presented distributions in position space as they are the easiest to conceptualize; however, SEPDA is capable of much more. Future studies should include the use of the momentum space distributions as well as the posmom intracule for each of the applications presented here, in addition to any others that could be fruitful.

In terms of the software development, there are some areas that could be pursued. Generally, both Fortran and C perform well in programs involving a plethora of calculations. The main calculation components involved in the SEPDA package are currently coded in Fortran. It could be useful to recode these components in

C to determine whether improvements in calculation time are obtained. For some calculations, especially those on large systems or those for $E(\boldsymbol{R})$ where multiple dimensions must be sampled, the time require can be rather prohibitive.

Another avenue to pursue in terms of decreasing time requirements is the implementation of a new algorithm to calculate the integrals involved in these electron pair distributions. Currently, Q-Chem[21] calculates intracules faster than does the SEPDA package which suggests that increases in speed could be obtained by using the same algorithms or by incorporating the SEPDA package into Q-Chem or a similar software package.

A final attempt to speed up these calculations could be made by converting the code to CUDA Fortran. CUDA is a parallel programming platform. Unpublished results by Hennessey and Klobukowski indicated a decrease in required calculation times by up to 18-fold on 2-electron integrals when going from Fortran code to CUDA Fortran.

The release of this software package to the general public should allow for SEPDA to be used in countless different applications that could not be achieved by the Pearson group alone. While various interesting applications were presented herein, the true value of this program will likely not be realized until its user base grows significantly.

Appendix A

SEPDA - User's Manual

This version was programmed by:

Adam J. Proud	University of Prince Edward Island

Contributions from:

Dalton Mackenzie	University of Prince Edward Island
Zosia Zielinski	University of Prince Edward Island
Brendan Sheppard	University of Prince Edward Island
Dr. Jason Pearson	University of Prince Edward Island
Dr. Joshua Hollett	University of Winnipeg
Dr. Peter Gill	Australia National University
Dr. Ajit Thakkar	University of New Brunswick

Contact information:

Dr. Jason Pearson
550 University Avenue
Charlottetown, PE C1A 4P3
jpearson@upei.ca

A.1 Introduction

This user's manual is intended to be used as a guide for users to understand the general use for the SEPDA software package (Single Electron Pair Distribution Analyzer), which was designed to allow for the analysis of electron pair behaviour within a single molecular orbital. This guide contains information regarding each of the parameters, both mandatory and optional, that are required for the proper performance of the program. The program must be run in a UNIX environment and users should have a general understanding of quantum chemical calculations.

A.1.1 Section Summaries

Section 1: Introduction to SEPDA and its capabilities

Section 2: Step by step instructions to install and run SEPDA on your local machine

Section 3: Input structure and variables

Section 4: Various examples of input files for different types of jobs

A.1.2 SEPDA Features

SEPDA 1.0 is designed to perform electron structure analyses on systems for which calculations have already been successfully performed. It is not a stand-alone program that can be used to perform the initial single point energy or geometry optimization calculations. The program takes molecular orbital coefficients and WFN files as input and subsequently calculates position/momentum/posmom intracules and/or extracules for a specific molecular orbital within the system.

The current version of SEPDA is capable of doing these calculations for localized molecular orbitals (LMOs from a GAMESS output), natural bond orbitals (NBOs from Weinhold's NBO program or any software package that utilizes said program), and canonical molecular orbitals (CMOs that can be calculated from any software package providing the relevant information is contained within the WFN file).

A.1.3 SEPDA Literature

The developers request that if you use the current software package that you cite the following articles. The appropriate literature will be indicated in all output files generated by SEPDA.

SEPDA General use:

Proud, A.J.; Sheppard, B.J.H.; Pearson, J.K. *J. Am. Chem. Soc.* **2018**, *140*, 219-228.

Scalar distributions:

Hollett, J.W.; Gill, P.M.W. J. Chem. Phys. 2011, 13, 2972-2978.

3-D distributions:

Thakkar, A.J. Moore, N.J. Int. J. Quantum Chem. 1981, 20, 393-400.

A.2 Installation and General Program Instructions

A.2.1 Installation Requirements

SEPDA is available for download at `https://bitbucket.org/aproud/sepda.git`. The package can be compiled using the install.sh script that is provided with the rest of the code. System requirements for installation are minimal but are as follows:

- An operating system with a UNIX-based command line
- Python version 2.x
- Fortran 90 compiler

A.2.2 Installing SEPDA

To install SEPDA, the user must simply run the install.sh script (./install.sh OR sh install.sh). This install script will define the SEPDADIR, SEPDABIN, and SEPDABASIS variables in each of the *.cshrc* and *.bashrc* files. Additionally, an alias will be defined in each of the aforementioned files so that *sepda* is recognized as a valid command. These variables will allow the user to run SEPDA from any directory. These variables define the following:

SEPDADIR	defines the location of the SEPDA directory. This is automatically defined as the path of the install script.
SEPDABIN	defines the location of the code to calculate the desired single electron pair density. The default is set to $SEPDA/bin and typically should not be changed
SEPDABASIS	defines the location of the basis set files. The default is set to $SEPDA/basis and typically should not be changed.
sepda	an alias to allow for the command *sepda* to be recognized as *$SEPDADIR/sepda.csh*

Should the user move the SEPDA directory following installation, these variables will need to be redefined in the appropriate shell profile files. To allow all users on a computer, these definitions could be incorporated into the system wide *.cshrc* and *.bashrc* files.

A.2.3 Running SEPDA

Following successful installation, the user should be able to run SEPDA from any directory using the sepda.csh script. Said script can be run in any of the following ways:

```
sepda infile.inp outfile.out
sepda infile.inp
```

In the event that an output file is not defined, the data from the calculation will be written to infile.dat. The structure of the infile.inp will be covered in detail in Chapter 3. Should the output file already exist, the original file will be overwritten with data from the new calculation.

The extensions of the infile and outfile are entirely up to the user and needn't be included. The following are two valid executions of the program:

```
sepda H2O.inp H2O.out
sepda water.inp
```

where in the first example, the input parameters should be defined in H2O.inp and the output will be directed to H2O.out, while in the second example, the input parameters are given in water.inp with the results contained within water.dat.

A.2.4 SEPDA Compatibility

As noted in Section 1.2, the current version of SEPDA is capable of performing calculations on 3 molecular orbital types: LMOs, CMOs, and NBOs.

A.2.5 Testing SEPDA

Within the $/SEPDADIR/examples directory, there are a variety of input files for testing to ensure that the program is functioning properly. Said directory also includes the ./output directory which contains the expected outputs from each of the sample inputs. These input files are those that will be included and described in detail in Chapter 4 of this manual and serve as a great reference for constructing your own input files.

A.3 Input File Structure and Variables

Input files for SEPDA require, at minimum, 3 input variables (WFN_FILE, BASIS, and MO_NUM); however, it is recommended that the user also specify the type of molecular orbital (MO_TYPE) as well as the type of pair density (PD_TYPE) to be calculated. In the event that the WFN file does not contain the molecular orbital coefficients, the COEF_FILE must also be specified. This would most commonly be the case when using NBOs as the MO_TYPE as the NBO coefficients with machine precision would be generated in the .37 file from the NBO program.

A.3.1 Description of Input Variables

A summary of each of the variables (both mandatory and optional) in the input file is given below along with their default value (if any).

Variable	**Default**	**Description**
WFN_FILE	None	defines the name of the .wfn file or file that contains the .wfn file
COEF_FILE	None	defines the name of the file containing the MO coefficients
BASIS	UD	states what basis set was used in the original calculation.
MO_TYPE	LMO	defines the type of MO to be analyzed. Options include LMO, CMO, and NBO
MO_NUM	None	defines what molecular orbital number on which to perform the desired calculation
PD_TYPE	1	defines the type of two-electron density to be calculated. Options include: 1 (position intracule), 2 (momentum intracule), 3 (posmom intracule), 4 (position extracule), 5 (momentum extracule), 6 (posmom extracule), 7 (3-D position extracule)
SCALE	Various	defines the scale parameter (α) for the Mura-Knowles grid. Smaller numbers lead to a more dense grid. Defaults are dependent on PD_TYPE
NPTS	Various	determines the number of points, i_{max}, to be calculated for the Mura-Knowles grid. Defaults are dependent on PD_TYPE

Additional details are required for a few of the input variables. This information is included in each of the subsections below.

WFN_FILE

The WFN file can be obtained from any software package capable of producing one. Programs such as GAMESS produce the contents of the WFN file within another file which is directed to the GAMESS scratch directory. Should the user be producing localized molecular orbital coefficients, they will be recorded in this same file. The user has the freedom to choose either the WFN file by itself or a file that contains the WFN file contents as the WFN_FILE variable.

COEF_FILE

In the case of both LMOs and CMOs, the coefficients (COEF_FILE) should be in the same file as the rest of the information to construct the wavefunction (WFN_FILE). For each of these cases, COEF_FILE needn't be specified. In the case that it is specified and is not the same as WFN_FILE the program will quit as it is assuming user error. For LMOs, both the .wfn file information and the LMO coefficients should be contained in the .dat file generated by GAMESS and redirected to the GAMESS scratch directory.

As for CMOs, all CMO coefficients are contained within the .wfn file, and thus, no other file is needed. To generate machine precision NBO coefficients in the atomic orbital basis, the user must use the following commands in either Gaussian or the NBO program:

```
$nbo aonbo=w37 file=nbocoefs $end
```

This will redirect the machine precision coefficients to the *nbocoefs.37* file.

BASIS

Note: The basis set in the input file must exactly match the one used for the initial calculation to obtain the molecular orbital coefficients, otherwise the calculation will fail.

The basis set must be defined in the input file. The user has the option to define their own basis set by setting BASIS equal to UD (for user-defined). If a user-defined basis set is desired, the input file must also include a basis section such as the following for a calculation on water:

```
$basis
   H     0
   S     2     1.00
      3.26800000     0.84570000
      0.98574000     0.23485000
   S     1     1.00
      0.24850000     1.00000000
   ####
   O     0
   S     3     1.00
      322.037000     0.05923940
      48.4308000     0.35150000
      10.4206000     0.70765800
   SP     2     1.00
      7.40294000     -0.4044530     0.24458600
      1.57620000     1.22156000     0.85395500
   SP 1 1.00
      0.37368400     1.00000000     1.00000000
   ####
$end
```

This is a standard basis set description used by most quantum chemical software packages. The first line denotes the atom, and all lines after describe the orbitals pertaining to that specific atom. One line must first designate the orbital type(s) (S, P, D, SP, or SPD), the level of contraction for said orbital, followed by 1.00. The spacing between these entries does

not matter as long as they are separated by at least one space. Depending on the level of contraction (n), the next n lines must describe the contracted Gaussian primitive (i.e. if the contraction level is 3, there must be 3 lines describing each Gaussian within that contracted set). On each of these n lines, there must be a Gaussian exponent, followed by a contraction coefficient. If the contraction level is just 1 (i.e. no contraction), said contraction coefficient should typically be set to unity. To begin the description of a new atom, one must add a break line that consists of four consecutive octothorpes (i.e. ####).

If the user wishes to use a pre-defined basis set, they should first verify that this basis set is available through the SEPDA software package. All available basis sets are included in the $SEPDABASIS directory as .bas files. As this code is only capable of calculating two-electron integrals for orbitals with angular momenta less than or equal to 2 (i.e. $l \leq 2$), any basis sets that have orbitals of higher angular momenta for the atoms in the given system will fail to run successfully. The user is free to add any .bas files to the $SEPDABASIS directory to simplify future calculations. These files must simply match the structure described above.

MO_NUM

The MO_NUM should be between 1 and either the number of Gaussian primitives (for NBOs) or the number of occupied orbitals (LMOs and CMOs). While the purpose of SEPDA is to calculate the two-electron densities of a single electron pair, setting MO_NUM equal to 0 leads to the calculation of the desired PD_TYPE for the full molecular system.

PD_TYPE

The PD_TYPE value should be between 1 and 7. The specific electron pair distribution corresponding to each value is provided in the table above describing each of the input variables. While PD_TYPE 7 is the 3D position extracule, only two dimensions (R_y and R_z

are actually sampled. This can be modified to calculate all 3 dimensions by commenting out line 681 and uncommenting line 682 in posvExt.f90. For both PD_TYPE 4 and 7, the user is advised to choose their coordinate system carefully. To adequately sample the Cartesian space around the molecular orbital of interest, the user must position the origin appropriately. The authors would suggest choosing an atom of interest as the origin for PD_TYPE 4, and either an atom or a bond midpoint as the origin for PD_TYPE 7. Furthermore, should the user wish to calculate a slice of the 3D position extracule (which is the default), it is advised that 3 atoms important to the MO of interest are utilized to define the yz-plane (that which is sampled by the grid).

SCALE and NPTS

Further control over calculations can be obtained by specifying the optional SCALE (α) and NPTS (i_{max}) variables. These options are used to define the grid values at which the value of the pair density will be determined and are related to the x-coordinate values by:

$$x = -\alpha \, ln \left[1 - \left(\frac{i}{i_{\text{max}}} \right)^3 \right], \qquad \text{where } i = 1, i_{\text{max}} \tag{A.1}$$

Each PD_TYPE has its own default values for SCALE and NPTS. It should be noted that for PD_TYPE 7 (i.e. 3-D position extracule), NPTS refers to the number of points in each direction of each Cartesian coordinate. Thus, if NPTS is set to 100, 100 grid points will be evaluated in the positive and negative directions of each Cartesian coordinate in addition to the origin. Therefore, 201^d (where $d = 1 - 3$, with the default being 2) grid points will be sampled for the calculation.

A.4 Examples of Input File Structures

ex1.inp

The input file below calculates a position intracule (PD_TYPE 1) density for a system (carbon atom) with the 3-21G basis set. The calculation is performed for canonical molecular orbital number 2. As neither SCALE nor NPTS are provided in the input file, the default values are used.

```
$init
          WFN_FILE     c_321G.dat
          BASIS        3-21G
          MO_TYPE      CMO
          PD_TYPE      1
          MO_NUM       2
$end
```

ex2.inp

The input file below calculates a momentum intracule (PD_TYPE 2) density for a system (methane) with the 6-31G basis set. The calculation is performed for natural bond orbital number 2. The default values for SCALE and NPTS are used as these variables were not specified in the input. Note that the COEFS variable needs to be defined as NBO coefficients are located in the .37 file. For instructions on how to generate the .37 file from a Gaussian or NBO output, please refer to section 3.1.2.

```
$init
          WFN_FILE     c_631G.wfn
          BASIS        6-31G
          MO_TYPE      NBO
          COEFS        CH4_631G.37
          PD_TYPE      2
          MO_NUM       2
$end
```

ex3.inp

The input file below calculates a posmom intracule (PD_TYPE 3) density for a system (carbon atom) using a user-defined (UD) basis set. Said basis set is described at the end of the input file in the $basis section according to the guidelines provided in section 3.1.3. The calculation is performed for canonical bond orbital number 3 using a grid with a SCALE value of 3.0 and 150 points.

```
$init
          WFN_FILE     c_UD.dat
          BASIS        UD
          MO_TYPE      CMO
          PD_TYPE      3
          MO_NUM       3
          SCALE        3.0
          NPTS         150
$end

$basis
   C     0
   S     3     1.00
      172.256000     0.06176690
      25.9109000     0.35879400
      5.53335000     0.70071300
   SP     2     1.00
      3.66498000     -0.3958970     0.23646000
      0.77054500     1.21584000     0.86061900
   SP 1 1.00
      0.19585700     1.00000000     1.00000000
   ####
$end
2
```

ex4.inp

The input file below calculates a position extracule (PD_TYPE 4) density for a system (carbon atom) with the 6-311G(d,p) basis set. The calculation is performed for localized molecular orbital number 2. Default values for both SCALE and NPTS are used since they were not specified.

```
$init
          WFN_FILE     c_6311GDP.dat
          BASIS        6-311G(d,p)
          MO_TYPE      LMO
          PD_TYPE      4
          MO_NUM       3
$end
```

ex5.inp

The input file below calculates a momentum extracule (PD_TYPE 5) density for a system (carbon atom) with the 3-21G basis set. The calculation is performed for canonical molecular orbital number 1. Default values for SCALE and NPTS are used as the user has not defined them in the input file.

```
$init
          WFN_FILE     c_321G.dat
          BASIS        3-21G
          MO_TYPE      CMO
          PD_TYPE      5
          MO_NUM       1
          SCALE        3.0
          NPTS         350
$end
```

ex6.inp

The input file below calculates a vectorized position extracule (PD_TYPE 7) density for a system (carbon atom) with the TZV basis set. This calculation is performed for localized molecular orbital number 3. It should be noted that NPTS 75 indicates that 75 points are sampled in each of the Cartesian coordinates and directions (i.e. 151 total for R_x, R_y, and R_z directions).

```
$init
          WFN_FILE     c_TZV.dat
          BASIS        TZV
          MO_TYPE      LMO
          PD_TYPE      7
          MO_NUM       3
          SCALE        2.1
          NPTS         75
$end
```

Appendix B

Derivation of Hollett and Gill Recurrence Relation

Many of the calculations involved in this thesis were performed using the recurrence relation developed by Hollett and Gill in 2010 or slightly modified versions for the purposes of extracule calculations. For this reason, a derivation of the RR will be provided here to show how the required integrals were obtained.

Boys demonstrated[26] that Gaussian basis functions of higher angular momenta could be determined from

$$|(\boldsymbol{a}+\mathbf{1_i})\rangle = \hat{D}|\boldsymbol{a}\rangle + \frac{a_i}{2\alpha}|(\boldsymbol{a}-\mathbf{1_i})\rangle \tag{B.1}$$

In this expression, $\hat{D}$ is the scaled differential operator which is defined as

$$\hat{D} = \frac{1}{2\alpha}\frac{\partial}{\partial A_i} \tag{B.2}$$

If the angular momentum augmentation is performed on any other basis function (i.e. b, c,

or d) the appropriate centres (B_i, C_i, or D_i) and exponents (β, γ, or δ) would replace A_i and α, respectively.

Before we proceed any further, we should introduce some important variables, some of which were presented in the text of this thesis but will be restated here for convenience:

$$\nu^2 = \frac{(\alpha+\beta)(\gamma+\delta)}{\alpha+\beta+\gamma+\delta} \tag{B.3}$$

$$\sigma^2 = \frac{1}{4\left(\frac{\alpha\beta}{\alpha+\beta}+\frac{\gamma\delta}{\gamma+\delta}\right)} \tag{B.4}$$

$$\lambda^2 = \frac{\alpha\delta}{\alpha+\delta}+\frac{\beta\gamma}{\beta+\gamma} \tag{B.5}$$

$$\eta = \frac{\alpha}{\alpha+\delta}-\frac{\beta}{\beta+\gamma} \tag{B.6}$$

$$\boldsymbol{U} = \frac{\alpha\boldsymbol{A}+\delta\boldsymbol{B}}{\alpha+\beta}-\frac{\gamma\boldsymbol{C}+\delta\boldsymbol{C}}{\gamma+\delta} \tag{B.7}$$

$$\boldsymbol{V} = \frac{2\alpha\beta}{\alpha+\beta}(\boldsymbol{A}-\boldsymbol{B})+\frac{2\gamma\delta}{\gamma+\delta}(\boldsymbol{D}-\boldsymbol{C}) \tag{B.8}$$

$$S_{ab} = \exp\left[-\frac{\alpha\beta|\boldsymbol{A}-\boldsymbol{B}|^2}{\alpha+\beta}-\frac{\gamma\delta|\boldsymbol{C}-\boldsymbol{D}|^2}{\gamma+\delta}\right] \tag{B.9}$$

To proceed, we must first define an operator $\hat{O}(\boldsymbol{a})$ that converts an s-type Gaussian primitive to one with an angular momentum, bsa. In other words

$$\hat{O}(\boldsymbol{a})|\boldsymbol{0}\rangle = |\boldsymbol{a}\rangle \tag{B.10}$$

This operator must also obey the Boys recurrence relation and thus

$$\hat{O}(\boldsymbol{a}+\boldsymbol{1_i}) = \hat{D}\hat{O}(\boldsymbol{a})+\frac{a_i}{2\alpha}\hat{O}(\boldsymbol{a}-\boldsymbol{1_i}) \tag{B.11}$$

Alrichs demonstrated that any function Y that is linear in A_i, has the commutation property as follows:

$$\hat{O}(a_i)Y = Y\hat{O}(a_i) + \frac{a_i y}{2\alpha}\hat{O}(a_i - 1) \tag{B.12}$$

where y is the partial derivative of Y with respect to the centre, A_i.

As the operators that will be considered ($\delta(u - r_{12})$ and $\delta(v - p_{12})$) are independent of the Gaussian centres ($\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\boldsymbol{D}$), any integral, $[\boldsymbol{abcd}]_Z$, can be determined as

$$[\boldsymbol{abcd}]_Z = \hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})[\boldsymbol{0000}]_Z \tag{B.13}$$

where Z is an arbitrary index denoting whatever type of integral is being dealt with at the time.

Assuming that we have the integral $[\boldsymbol{abcd}]_Z$, by applying the Boys RR described in equation (B.1), we obtain

$$[(\boldsymbol{a} + \boldsymbol{1_i})\boldsymbol{bcd}]_Z = \hat{D}[\boldsymbol{abcd}]_Z + \frac{a_i}{2\alpha}[(\boldsymbol{a} - \boldsymbol{1_i})\boldsymbol{bcd}]_Z \tag{B.14}$$

However, to obtain this integral, we must first apply equation (B.13). Applying this equation to (B.14) yields

$$[(\boldsymbol{a} + \boldsymbol{1_i})\boldsymbol{bcd}]_Z = \hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\hat{D}[\boldsymbol{0000}]_Z + \frac{a_i}{2\alpha}[(\boldsymbol{a} - \boldsymbol{1_i})\boldsymbol{bcd}]_Z \tag{B.15}$$

This discussion has thus far has been completely general. Recall from Chapter 1 that both position and momentum intracules and extracules can be determined from the second-order reduced Wigner distribution, $W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2})$. A general expression can be written for the determination of intracules or extracules by assuming a general two-electron phase-

space operator, $\hat{L}$. Then, the desired electron pair distribution can be obtained as the "expectation value" of the operator as

$$\langle \hat{L} \rangle = \int W(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) \hat{L}(\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{p_1}, \boldsymbol{p_2}) \mathrm{d}\boldsymbol{r_1}\, \mathrm{d}\boldsymbol{r_2}\, \mathrm{d}\boldsymbol{p_1}\, \mathrm{d}\boldsymbol{p_2} \tag{B.16}$$

For the purposes of this derivation, let us consider the position and momentum intracules, $P(u)$ and $M(v)$. For these two cases, $\hat{L}$ is defined as $\delta(u - r_{12})$ and $\delta(v - p_{12})$, respectively. Regardless, the derivation is largely the same.

The fundamental integral, $[\mathbf{0000}]_W$, for these calculations is given by

$$\begin{aligned}[\mathbf{0000}]_W =& \frac{S_{ab} e^{-\lambda^2 U^2 + \mu^2 V^2 + \eta \boldsymbol{U}\cdot\boldsymbol{V}}}{8(\alpha+\delta)^{3/2}(\beta+\gamma)^{3/2}} \\ & \int e^{-\lambda^2 u^2 - \mu^2 \nu^2 - (2\lambda^2 \boldsymbol{U} - \eta \boldsymbol{V})\cdot\boldsymbol{u} - i(\eta \boldsymbol{U} + 2\mu^2 \boldsymbol{V})\cdot\boldsymbol{v}} \hat{L}(u,v) \mathrm{d}\boldsymbol{u}\, \mathrm{d}\boldsymbol{v}\end{aligned} \tag{B.17}$$

When differentiating this expression, four terms will result due to the dependence of S_{ab}, U^2, V^2, and $\boldsymbol{U}\cdot\boldsymbol{V}$, on the centre $\boldsymbol{A}$. For this reason, it is useful to define the following function

$$\begin{aligned}G_{l,m,n}(U^2, V^2, \boldsymbol{U}\cdot\boldsymbol{V}) =& \frac{\left(\frac{\partial}{\partial U^2}\right)^l \left(\frac{\partial}{\partial V^2}\right)^m \left(\frac{\partial}{\partial(\boldsymbol{U}\cdot\boldsymbol{V})}\right)^n}{8(\alpha+\delta)^{3/2}(\beta+\gamma)^{3/2}} e^{-\lambda^2 U^2 + \mu^2 V^2 + \eta \boldsymbol{U}\cdot\boldsymbol{V}} \\ & \int e^{-\lambda^2 u^2 - \mu^2 \nu^2 - (2\lambda^2 \boldsymbol{U} - \eta \boldsymbol{V})\cdot\boldsymbol{u} - i(\eta \boldsymbol{U} + 2\mu^2 \boldsymbol{V})\cdot\boldsymbol{v}} \hat{L}(u,v) \mathrm{d}\boldsymbol{u}\, \mathrm{d}\boldsymbol{v}\end{aligned} \tag{B.18}$$

in addition to the triple-index integrals

$$[\mathbf{0000}]_W^{(l,m,n)} = S_{ab} G_{l,m,n}(U^2, V^2, \boldsymbol{U}\cdot\boldsymbol{V}) \tag{B.19}$$

Taking this expression and substituting it into equation (B.15) gives

$$[(\boldsymbol{a}+\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)} = \hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\hat{D}(S_{ab}G_{l,m,n}(U^2,V^2,\boldsymbol{U}\cdot\boldsymbol{V})) + \frac{a_i}{2\alpha}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)} \tag{B.20}$$

and applying the scaled differential operator yields

$$\begin{aligned}[(\boldsymbol{a}+\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)} =& \hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\mathbf{0000}]_W^{(l,m,n)}\right.\\ &+\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\mathbf{0000}]_W^{(l+1,m,n)}+\frac{1}{2\alpha}\frac{\partial V^2}{\partial A_i}[\mathbf{0000}]_W^{(l,m+1,n)}\\ &\left.+\frac{1}{2\alpha}\frac{\partial(\boldsymbol{U}\cdot\boldsymbol{V}}{\partial A_i}[\mathbf{0000}]_W^{(l,m,n+1)}\right)+\frac{a_i}{2\alpha}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)}\end{aligned} \tag{B.21}$$

To more clearly demonstrate the remainder of the derivation, the terms in this expression have been color coded based on different values for the triple indices, (l,m,n). Each term will be dealt with separately. Applying the commutation property from equation (B.12) to the (l,m,n) integrals (those in blue) four times for each operator, $\hat{O}(\boldsymbol{x})$, gives:

$$
\begin{aligned}
&\hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{0000}]_W^{(l,m,n)}\right)= \\
&\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{a000}]_W^{(l,m,n)}-\frac{a_i}{2\alpha}\frac{\beta}{\alpha+\beta}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{000}]_W^{(l,m,n)}\right)= \\
&\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{ab00}]_W^{(l,m,n)}-\frac{a_i}{2\alpha}\frac{\beta}{\alpha+\beta}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{b00}]_W^{(l,m,n)}\right. \\
&\qquad\left.+\frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\boldsymbol{1_i})\boldsymbol{00}]_W^{(l,m,n)}\right)= \\
&\hat{O}(\boldsymbol{d})\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{abc0}]_W^{(l,m,n)}-\frac{a_i}{2\alpha}\frac{\beta}{\alpha+\beta}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{bc0}]_W^{(l,m,n)}\right. \\
&\qquad\left.+\frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\boldsymbol{1_i})\boldsymbol{c0}]_W^{(l,m,n)}\right)= \\
&\left(\frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l,m,n)}-\frac{a_i}{2\alpha}\frac{\beta}{\alpha+\beta}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{bcd}]_W^{(l,m,n)}\right. \\
&\qquad\left.+\frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\boldsymbol{1_i})\boldsymbol{cd}]_W^{(l,m,n)}\right) \qquad \text{(B.22)}
\end{aligned}
$$

The observant reader may notice that no new terms were introduced after operation of $\hat{O}(\boldsymbol{c})$ and $\hat{O}(\boldsymbol{d})$. This is caused by the fact that the original term is independent of both centres, $\boldsymbol{C}$ and $\boldsymbol{D}$. Thus, the derivative, y, is equal to zero and the term vanishes. If we now then do this same procedure to the integrals highlighted in green , $(l+1,m,n)$, we obtain:

$$
\begin{aligned}
&\hat{O}(\boldsymbol{a})\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\boldsymbol{0000}]_W^{(l+1,m,n)}\right)= \\
&\hat{O}(\boldsymbol{b})\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\boldsymbol{a000}]_W^{(l+1,m,n)}+\frac{a_i}{4\alpha^2}\frac{\partial^2 U^2}{\partial A_i^2}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{000}]_W^{(l+1,m,n)}\right)= \\
&\hat{O}(\boldsymbol{c})\hat{O}(\boldsymbol{d})\left(\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\boldsymbol{ab00}]_W^{(l+1,m,n)}+\frac{a_i}{4\alpha^2}\frac{\partial^2 U^2}{\partial A_i^2}[(\boldsymbol{a}-\boldsymbol{1_i})\boldsymbol{b00}]_W^{(l+1,m,n)}\right. \\
&\qquad\left.+\frac{b_i}{4\alpha\beta}\frac{\partial^2 U^2}{\partial A_i\partial B_i}[\boldsymbol{a}(\boldsymbol{b}-\boldsymbol{1_i})\boldsymbol{00}]_W^{(l+1,m,n)}\right)=
\end{aligned}
$$

$$\hat{O}(\boldsymbol{d})\left(\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\boldsymbol{abc0}]_W^{(l+1,m,n)}+\frac{a_i}{4\alpha^2}\frac{\partial^2 U^2}{\partial A_i^2}[\boldsymbol{(a-1_i)bc0}]_W^{(l+1,m,n)}\right.$$
$$\left.+\frac{b_i}{4\alpha\beta}\frac{\partial^2 U^2}{\partial A_i\partial B_i}[\boldsymbol{a(b-1_i)c0}]_W^{(l+1,m,n)}+\frac{c_i}{4\alpha\gamma}\frac{\partial^2 U^2}{\partial A_i\partial C_i}[\boldsymbol{ab(c-1_i)0}]_W^{(l+1,m,n)}\right)=$$
$$\left(\frac{1}{2\alpha}\frac{\partial U^2}{\partial A_i}[\boldsymbol{abcd}]_W^{(l+1,m,n)}+\frac{a_i}{4\alpha^2}\frac{\partial^2 U^2}{\partial A_i^2}[\boldsymbol{(a-1_i)bcd}]_W^{(l+1,m,n)}\right.$$
$$+\frac{b_i}{4\alpha\beta}\frac{\partial^2 U^2}{\partial A_i\partial B_i}[\boldsymbol{a(b-1_i)cd}]_W^{(l+1,m,n)}+\frac{c_i}{4\alpha\gamma}\frac{\partial^2 U^2}{\partial A_i\partial C_i}[\boldsymbol{ab(c-1_i)d}]_W^{(l+1,m,n)}$$
$$\left.+\frac{d_i}{4\alpha\delta}\frac{\partial^2 U^2}{\partial A_i\partial D_i}[\boldsymbol{abc(d-1_i)}]_W^{(l+1,m,n)}\right) \tag{B.23}$$

In this final expression, there exists multiple partial derivatives that need be evaluated. These derivatives are equal to the following:

$$\frac{\partial U^2}{\partial A_i}=2U_i\frac{\partial U_i}{\partial A_i}=2U_i\frac{\alpha}{\alpha+\beta} \tag{B.24}$$
$$\frac{\partial^2 U^2}{\partial A_i^2}=\frac{\partial}{\partial A_i}\frac{\partial U^2}{\partial A_i}=\frac{2\alpha}{\alpha+\beta}\frac{\alpha}{\alpha+\beta}=\frac{2\alpha^2}{(\alpha+\beta)^2} \tag{B.25}$$
$$\frac{\partial^2 U^2}{\partial A_i\partial B_i}=\frac{\partial}{\partial B_i}\frac{\partial U^2}{\partial A_i}=\frac{2\beta}{\alpha+\beta}\frac{\alpha}{\alpha+\beta}=\frac{2\alpha\beta}{(\alpha+\beta)^2} \tag{B.26}$$
$$\frac{\partial^2 U^2}{\partial A_i\partial C_i}=\frac{\partial}{\partial C_i}\frac{\partial U^2}{\partial A_i}=-\frac{2\gamma}{\gamma+\delta}\frac{\alpha}{\alpha+\beta}=-\frac{2\alpha\gamma}{(\alpha+\beta)(\gamma+\delta)} \tag{B.27}$$
$$\frac{\partial^2 U^2}{\partial A_i\partial D_i}=\frac{\partial}{\partial C_i}\frac{\partial U^2}{\partial A_i}=-\frac{2\delta}{\gamma+\delta}\frac{\alpha}{\alpha+\beta}=-\frac{2\alpha\delta}{(\alpha+\beta)(\gamma+\delta)} \tag{B.28}$$

Substituting these expressions into equation (B.23) gives a final expression for the $(l+1,m,n)$ integrals:

$$\left(\frac{U_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l+1,m,n)}+\frac{a_i}{2(\alpha+\beta)^2}[\boldsymbol{(a-1_i)bcd}]_W^{(l+1,m,n)}+\frac{b_i}{2(\alpha+\beta)^2}[\boldsymbol{a(b-1_i)cd}]_W^{(l+1,m,n)}\right.$$
$$\left.-\frac{c_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab(c-1_i)d}]_W^{(l+1,m,n)}-\frac{d_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc(d-1_i)}]_W^{(l+1,m,n)}\right) \tag{B.29}$$

As both the $(l,m+1,n)$ and $(l,m,n+1)$ integrals are calculated much the same way as

the $(l, m, n+1)$ integrals, the complete derivation will not be shown. The final expressions for these integrals are

$$\left(\frac{2\beta V_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l,m+1,n)} + \frac{2a_i\beta^2}{(\alpha+\beta)^2}[(\boldsymbol{a-1_i})\boldsymbol{bcd}]_W^{(l,m+1,n)} + \frac{2b_i\alpha\beta}{(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b-1_i})\boldsymbol{cd}]_W^{(l,m+1,n)} - \frac{2c_i\beta\delta}{(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c-1_i})\boldsymbol{d}]_W^{(l,m+1,n)} - \frac{2d_i\beta\gamma}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d-1_i})]_W^{(l,m+1,n)}\right) \tag{B.30}$$

and

$$\left(\frac{2\beta U_i + V_i}{2(\alpha+\beta)}[\boldsymbol{abcd}]_W^{(l,m,n+1)} + \frac{a_i\beta}{(\alpha+\beta)^2}[(\boldsymbol{a-1_i})\boldsymbol{bcd}]_W^{(l,m,n+1)} + \frac{b_i(\beta-\alpha)}{2(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b-1_i})\boldsymbol{cd}]_W^{(l,m,n+1)} - \frac{c_i(\beta+\delta)}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c-1_i})\boldsymbol{d}]_W^{(l,m,n+1)} - \frac{d_i(\beta-\gamma)}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d-1_i})]_W^{(l,m,n+1)}\right) \tag{B.31}$$

Taking each of these final expressions and plugging them into equation (B.21) yields a 19 term expression that can be simplified to the following 18 terms by combining the term outside the parentheses in (B.21) with a similar term in (B.22). The final result is

$$
\begin{aligned}
[(\boldsymbol{a}+\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)} = \Bigg(& \frac{\beta(B_i-A_i)}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l,m,n)} + \frac{a_i}{2(\alpha+\beta)}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n)} \\
& + \frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\mathbf{00}]_W^{(l,m,n)} + \frac{U_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l+1,m,n)} + \frac{a_i}{2(\alpha+\beta)^2}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l+1,m,n)} \\
& + \frac{b_i}{2(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\boldsymbol{cd}]_W^{(l+1,m,n)} - \frac{c_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c}-\mathbf{1_i})\boldsymbol{d}]_W^{(l+1,m,n)} \\
& - \frac{d_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d}-\mathbf{1_i})]_W^{(l+1,m,n)} + \frac{2\beta V_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l,m+1,n)} \\
& + \frac{2a_i\beta^2}{(\alpha+\beta)^2}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m+1,n)} + \frac{2b_i\alpha\beta}{(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\boldsymbol{cd}]_W^{(l,m+1,n)} \\
& - \frac{2c_i\beta\delta}{(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c}-\mathbf{1_i})\boldsymbol{d}]_W^{(l,m+1,n)} - \frac{2d_i\beta\gamma}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d}-\mathbf{1_i})]_W^{(l,m+1,n)} \\
& + \frac{2\beta U_i+V_i}{2(\alpha+\beta)}[\boldsymbol{abcd}]_W^{(l,m,n+1)} + \frac{a_i\beta}{(\alpha+\beta)^2}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l,m,n+1)} \\
& + \frac{b_i(\beta-\alpha)}{2(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\boldsymbol{cd}]_W^{(l,m,n+1)} - \frac{c_i(\beta+\delta)}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c}-\mathbf{1_i})\boldsymbol{d}]_W^{(l,m,n+1)} \\
& - \frac{d_i(\beta-\gamma)}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d}-\mathbf{1_i})]_W^{(l,m,n+1)} \Bigg)
\end{aligned}
\tag{B.32}
$$

In the cases where the operator $\hat{L}$ depends only on $\boldsymbol{u}$ or $\boldsymbol{v}$, $G_{l,m,n}(U^2,V^2,\boldsymbol{U}\cdot\boldsymbol{V})$ simplifies to

$$
G_{l,m,n}(U^2,V^2,\boldsymbol{U}\cdot\boldsymbol{V}) = \frac{\pi^{3/2}\left(\frac{\partial}{\partial U^2}\right)^l\left(\frac{\partial}{\partial V^2}\right)^m\left(\frac{\partial}{\partial(\boldsymbol{U}\cdot\boldsymbol{V})}\right)^n}{(\alpha+\beta+\gamma+\delta)^{3/2}}\int e^{-\nu^2|\boldsymbol{u}+\boldsymbol{U}|^2}\hat{L}(\boldsymbol{u})\mathrm{d}\boldsymbol{u} \tag{B.33}
$$

$$
G_{l,m,n}(U^2,V^2,\boldsymbol{U}\cdot\boldsymbol{V}) = \frac{\pi^{3/2}\sigma^3\left(\frac{\partial}{\partial U^2}\right)^l\left(\frac{\partial}{\partial V^2}\right)^m\left(\frac{\partial}{\partial(\boldsymbol{U}\cdot\boldsymbol{V})}\right)^n}{(\alpha+\beta)^{3/2}(\gamma+\delta)^{3/2}}\int e^{-\sigma^2|\boldsymbol{v}+i\boldsymbol{V}|^2}\hat{L}(\boldsymbol{v})\mathrm{d}\boldsymbol{v} \tag{B.34}
$$

More importantly, it can be shown that for these two special cases:

$$G_{l,m,n}(U^2, V^2, \boldsymbol{U}\cdot\boldsymbol{V}) = 0 \qquad \text{for } \hat{L}(\boldsymbol{u},\boldsymbol{v}) = \hat{L}(\boldsymbol{u}) \text{ if } m > 0 \text{ or } n > 0 \tag{B.35}$$

$$G_{l,m,n}(U^2, V^2, \boldsymbol{U}\cdot\boldsymbol{V}) = 0 \qquad \text{for } \hat{L}(\boldsymbol{u},\boldsymbol{v}) = \hat{L}(\boldsymbol{v}) \text{ if } l > 0 \text{ or } n > 0 \tag{B.36}$$

Therefore, in the case of the position intracule, $P(u)$, where $\hat{L} = \delta(u - r_{12})$, the 18-term recurrence relation reduces to the following 8 terms:

$$\begin{aligned}
[(\boldsymbol{a}+\mathbf{1_i})\boldsymbol{bcd}]_W^{(l)} = \Bigg(& \frac{\beta(B_i - A_i)}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l)} + \frac{a_i}{2(\alpha+\beta)}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l)} \\
& + \frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\mathbf{00}]_W^{(l)} + \frac{U_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(l+1)} + \frac{a_i}{2(\alpha+\beta)^2}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(l+1)} \\
& + \frac{b_i}{2(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\boldsymbol{cd}]_W^{(l+1)} - \frac{c_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c}-\mathbf{1_i})\boldsymbol{d}]_W^{(l+1)} \\
& - \frac{d_i}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d}-\mathbf{1_i})]_W^{(l+1)} \Bigg)
\end{aligned} \tag{B.37}$$

where the m and n indices have been removed as they are no longer required. Similarly, with this same operator, $G_{l,m,n}$ becomes

$$G_l(U^2) = \frac{4\pi^{5/2}u^2 e^{-\nu^2 u^2}}{(\alpha+\beta+\gamma+\delta)^{3/2}} \left(\frac{\partial}{\partial U^2}\right)^l [e^{-\nu^2 U^2} i_0(2\nu^2) U u] \tag{B.38}$$

where

$$i_0(x) = \frac{\sinh x}{x} \tag{B.39}$$

Similarly, for the momentum intracule, $M(v)$, where $\hat{L} = \delta(v - p_{12})$, the recurrence relation simplifies to

$$
\begin{aligned}
[(\boldsymbol{a}+\mathbf{1_i})\boldsymbol{bcd}]_W^{(m)} = \Bigg(& \frac{\beta(B_i - A_i)}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(m)} + \frac{a_i}{2(\alpha+\beta)}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(m)} \\
& + \frac{b_i}{2(\alpha+\beta)}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\mathbf{00}]_W^{(m)} + \frac{2\beta V_i}{\alpha+\beta}[\boldsymbol{abcd}]_W^{(m+1)} \\
& + \frac{2a_i\beta^2}{(\alpha+\beta)^2}[(\boldsymbol{a}-\mathbf{1_i})\boldsymbol{bcd}]_W^{(m+1)} + \frac{2b_i\alpha\beta}{(\alpha+\beta)^2}[\boldsymbol{a}(\boldsymbol{b}-\mathbf{1_i})\boldsymbol{cd}]_W^{(m+1)} \\
& - \frac{2c_i\beta\delta}{(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{ab}(\boldsymbol{c}-\mathbf{1_i})\boldsymbol{d}]_W^{(m+1)} - \frac{2d_i\beta\gamma}{2(\alpha+\beta)(\gamma+\delta)}[\boldsymbol{abc}(\boldsymbol{d}-\mathbf{1_i})]_W^{(m+1)} \Bigg)
\end{aligned}
\tag{B.40}
$$

In this case, the l and n indices are dropped as values over 0 cause the function to vanish. As before, $G_{l,m,n}$, simplifies again to

$$
G_m(V^2) = \frac{4\pi^{5/2}\sigma^3 v^2 e^{-\sigma^2 u^2}}{(\alpha+\beta)^{3/2}(\gamma+\delta)^{3/2}} \left(\frac{\partial}{\partial V^2}\right)^m [e^{\sigma^2 V^2} j_0(2\sigma^2) V v] \tag{B.41}
$$

where

$$
j_0(x) = \frac{\sin x}{x} \tag{B.42}
$$

Using the recurrence relations described by equations (B.37) and (B.40) combined with the definition of the fundamental integral in (B.19), $P(u)$ and $M(v)$ can be calculated rather efficiently. Similar derivations could be performed for the position and momentum extracules, $E(R)$ and $E(P)$.

Appendix C

Supplementary Information for the Analysis of Hydrogen Bonding Complexes

This section consists of the tables containing all of the data for the metrics for the hydrogen bonding analyses presented in Chapter 6. Data is presented for both $E(\boldsymbol{R})$ and $P(u)$ for the $\sigma_{\text{X-H}}$ bond LMOs and the n_{Y} lone pairs at all distances previously discussed. Atomic units are used throughout this chapter.

Table C.1: $\delta_{\boldsymbol{R}}$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$
$AcNH_2$-$AcNH_2$	0.0964	0.0816	0.0695	0.0598	0.0513	0.0341	0.0192	0.0064
AcOH-AcOH	0.1262	0.1065	0.0902	0.0766	0.0653	0.0414	0.0211	0.0068
CCl_3OH-H_2O	0.1050	0.0886	0.0753	0.0642	0.0400	0.0357	0.0193	0.0074
CF_3OH-H_2O	0.0928	0.0783	0.0664	0.0567	0.0486	0.0317	0.0174	0.0068
HBr-MeOH	0.0761	0.0641	0.0546	0.0468	0.0404	0.0271	0.0158	0.0067
HCl-$MeNH_2$	0.1655	0.1418	0.1216	0.1047	0.0906	0.0608	0.0338	0.0131
HCl-MeOH	0.0798	0.0672	0.0571	0.0488	0.0421	0.0421	0.0162	0.0068
HF-$MeNH_2$	0.1432	0.1225	0.1053	0.0909	0.0787	0.0513	0.0265	0.0095
HF-MeOH	0.0937	0.0789	0.0666	0.0563	0.0478	0.0302	0.0165	0.0066
$MeNH_2$-$MeNH_2$	0.0550	0.0462	0.0390	0.0331	0.0283	0.0185	0.0087	0.0030
$MeNH_2$-MeOH	0.0427	0.0355	0.0298	0.0252	0.0216	0.0143	0.0078	0.0031
$MeNH_2$-Peptide	0.0491	0.0415	0.0354	0.0305	0.0264	0.0150	0.0067	0.0025
$MeNH_2$-Pyridine	0.0369	0.0305	0.0253	0.0212	0.0179	0.0116	0.0065	0.0026
MeOH-CH_3Cl	0.0326	0.0252	0.0198	0.0157	0.0126	0.0074	0.0029	0.0009
MeOH-CH_3F	0.0397	0.0321	0.0262	0.0216	0.0181	0.0114	0.0058	0.0023
MeOH-$MeNH_2$	0.0978	0.0833	0.0713	0.0611	0.0526	0.0343	0.0186	0.0067
MeOH-MeOH	0.0727	0.0607	0.0509	0.0430	0.0364	0.0232	0.0125	0.0046
MeOH-Peptide	0.0919	0.0782	0.0669	0.0574	0.0496	0.0331	0.0173	0.0048
MeOH-Pyridine	0.0939	0.0794	0.0674	0.0580	0.0494	0.0313	0.0169	0.0063
MeOH-H_2O	0.0656	0.0548	0.0461	0.0391	0.0334	0.0218	0.0121	0.0047
Peptide-$MeNH_2$	0.0778	0.0660	0.0563	0.0483	0.0416	0.0273	0.0150	0.0055
Peptide-MeOH	0.0566	0.0474	0.0400	0.0341	0.0293	0.0195	0.0112	0.0045
Peptide-H_2O	0.0511	0.0428	0.0363	0.0311	0.0270	0.0185	0.0148	0.0046
H_2O-$MeNH_2$	0.0873	0.0741	0.0631	0.0540	0.0463	0.0301	0.0163	0.0061
H_2O-MeOH	0.0681	0.0567	0.0474	0.0399	0.0338	0.0214	0.0115	0.0043
H_2O-Peptide	0.1205	0.1114	0.1041	0.0985	0.0942	0.0864	0.0813	0.0797
H_2O-Pyridine	0.1161	0.1062	0.0987	0.0931	0.0889	0.0815	0.0775	0.0763
H_2O-H_2O	0.0614	0.0512	0.0430	0.0364	0.0310	0.0202	0.0112	0.0044

Table C.2: $\langle \boldsymbol{R}^0_{\boldsymbol{yz}} \rangle$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.8452	0.8435	0.8418	0.8410	0.8391	0.8365	0.8344	0.8323	0.8309
AcOH-AcOH	0.9626	0.9603	0.9583	0.9565	0.9548	0.9510	0.9479	0.9458	0.9447
CCl_3OH-H_2O	0.9598	0.9581	0.9565	0.9550	0.9516	0.9510	0.9489	0.9475	0.9465
CF_3OH-H_2O	0.9608	0.9591	0.9576	0.9562	0.9551	0.9527	0.9509	0.9494	0.9485
HBr-MeOH	0.6621	0.6632	0.6640	0.6646	0.6651	0.6663	0.6675	0.6679	0.6683
HCl-$MeNH_2$	0.7174	0.7192	0.7206	0.7217	0.7227	0.7248	0.7262	0.7272	0.7271
HCl-MeOH	0.7340	0.7347	0.7351	0.7354	0.7356	0.7356	0.7369	0.7369	0.7369
HF-$MeNH_2$	1.0749	1.0712	1.0681	1.0655	1.0631	1.0576	1.0526	1.0498	1.0480
HF-MeOH	1.0724	1.0697	1.0673	1.0652	1.0634	1.0597	1.0572	1.0553	1.0541
$MeNH_2$-$MeNH_2$	0.8418	0.8408	0.8401	0.8396	0.8392	0.8385	0.8380	0.8374	0.8371
$MeNH_2$-MeOH	0.8414	0.8407	0.8401	0.8397	0.8395	0.8390	0.8384	0.8379	0.8375
$MeNH_2$-Peptide	0.8427	0.8419	0.8413	0.8409	0.8405	0.8396	0.8385	0.8380	0.8377
$MeNH_2$-Pyridine	0.8406	0.8399	0.8394	0.8390	0.8387	0.8383	0.8380	0.8376	0.8373
MeOH-CH_3Cl	0.9512	0.9498	0.9487	0.9480	0.9475	0.9468	0.9461	0.9458	0.9457
MeOH-CH_3F	0.9521	0.9509	0.9499	0.9492	0.9487	0.9478	0.9468	0.9462	0.9459
MeOH-$MeNH_2$	0.9580	0.9558	0.9539	0.9521	0.9507	0.9475	0.9451	0.9432	0.9422
MeOH-MeOH	0.9471	0.9453	0.9438	0.9426	0.9416	0.9396	0.9383	0.9372	0.9366
MeOH-Peptide	0.9488	0.9470	0.9453	0.9438	0.9426	0.9401	0.9380	0.9361	0.9354
MeOH-Pyridine	0.9569	0.9550	0.9532	0.9523	0.9507	0.9475	0.9455	0.9436	0.9427
MeOH-H_2O	0.9489	0.9472	0.9458	0.9446	0.9437	0.9419	0.9407	0.9395	0.9389
Peptide-$MeNH_2$	0.8377	0.8371	0.8365	0.8360	0.8355	0.8347	0.8344	0.8339	0.8336
Peptide-MeOH	0.8387	0.8381	0.8377	0.8373	0.8371	0.8367	0.8365	0.8361	0.8358
Peptide-H_2O	0.8425	0.8432	0.8428	0.8425	0.8422	0.8417	0.8404	0.8404	0.8397
H_2O-$MeNH_2$	0.9529	0.9509	0.9490	0.9475	0.9462	0.9436	0.9418	0.9403	0.9394
H_2O-MeOH	0.9481	0.9464	0.9450	0.9438	0.9429	0.9412	0.9400	0.9389	0.9383
H_2O-Peptide	0.9523	0.9503	0.9487	0.9473	0.9462	0.9440	0.9425	0.9408	0.9443
H_2O-Pyridine	0.9524	0.9505	0.9488	0.9474	0.9462	0.9439	0.9423	0.9409	0.9444
H_2O-H_2O	0.9514	0.9498	0.9484	0.9473	0.9465	0.9449	0.9438	0.9427	0.9421

Table C.3: $R_z^{\max}$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.0278	0.0556	0.0818	0.1008	0.1124	0.1330	0.1483	0.1598	0.1649
AcOH-AcOH	-0.2237	-0.2154	-0.2082	-0.2021	-0.1971	-0.1863	-0.1757	-0.1672	-0.1629
CCl_3OH-H_2O	-0.2049	-0.1976	-0.1915	-0.1864	-0.1755	-0.1734	-0.1644	-0.1571	-0.1522
CF_3OH-H_2O	-0.1866	-0.1797	-0.1741	-0.1696	-0.1658	-0.1574	-0.1490	-0.1421	-0.1373
HBr-MeOH	0.5401	0.5539	0.5641	0.5715	0.5778	0.5906	0.6020	0.6107	0.6176
HCl-$MeNH_2$	0.1049	0.1208	0.1361	0.1521	0.1713	0.3324	0.3788	0.4052	0.4202
HCl-MeOH	0.3241	0.3488	0.3641	0.3748	0.3829	0.3829	0.4108	0.4207	0.4275
HF-$MeNH_2$	-0.3171	-0.3138	-0.3109	-0.3085	-0.3064	-0.3021	-0.2980	-0.2947	-0.2928
HF-MeOH	-0.2898	-0.2873	-0.2853	-0.2838	-0.2825	-0.2797	-0.2771	-0.2751	-0.2737
$MeNH_2$-$MeNH_2$	0.1869	0.1919	0.1958	0.1989	0.2015	0.2072	0.2132	0.2165	0.2182
$MeNH_2$-MeOH	0.1948	0.1985	0.2016	0.2040	0.2060	0.2103	0.2140	0.2167	0.2186
$MeNH_2$-Peptide	0.1912	0.1951	0.1982	0.2008	0.2030	0.2097	0.2147	0.2172	0.2187
$MeNH_2$-Pyridine	0.1994	0.2024	0.2048	0.2068	0.2084	0.2119	0.2149	0.2171	0.2187
MeOH-CH_3Cl	-0.1010	-0.0969	-0.0937	-0.0911	-0.0889	-0.0849	-0.0811	-0.0793	-0.0785
MeOH-CH_3F	-0.1061	-0.1017	-0.0982	-0.0952	-0.0927	-0.0876	-0.0832	-0.0801	-0.0781
MeOH-$MeNH_2$	-0.1564	-0.1478	-0.1407	-0.1348	-0.1298	-0.1181	-0.1062	-0.0967	-0.0907
MeOH-MeOH	-0.1318	-0.1248	-0.1190	-0.1142	-0.1100	-0.1007	-0.0922	-0.0858	-0.0816
MeOH-Peptide	-0.1494	-0.1416	-0.1351	-0.1297	-0.1250	-0.1141	-0.1017	-0.0910	-0.0865
MeOH-Pyridine	-0.1527	-0.1439	-0.1367	-0.1307	-0.1257	-0.1143	-0.1033	-0.0947	-0.0891
MeOH-H_2O	-0.1255	-0.1193	-0.1143	-0.1100	-0.1064	-0.0982	-0.0906	-0.0846	-0.0804
Peptide-$MeNH_2$	0.0987	0.1155	0.1268	0.1351	0.1414	0.1539	0.1644	0.1720	0.1762
Peptide-MeOH	0.1318	0.1400	0.1460	0.1507	0.1544	0.1622	0.1690	0.1742	0.1776
Peptide-H_2O	0.1460	0.1522	0.1572	0.1610	0.1641	0.1705	0.1766	0.1808	0.1841
H_2O-$MeNH_2$	-0.1465	-0.1384	-0.1318	-0.1263	-0.1216	-0.1109	-0.1003	-0.0919	-0.0864
H_2O-MeOH	-0.1273	-0.1205	-0.1150	-0.1104	-0.1064	-0.0976	-0.0897	-0.0838	-0.0798
H_2O-Peptide	-0.1433	-0.1358	-0.1296	-0.1244	-0.1200	-0.1097	-0.0994	-0.0900	-0.0708
H_2O-Pyridine	-0.1440	-0.1358	-0.1291	-0.1236	-0.1190	-0.1086	-0.0988	-0.0912	-0.0711
H_2O-H_2O	-0.1213	-0.1154	-0.1106	-0.1065	-0.1031	-0.0954	-0.0884	-0.0828	-0.0789

Table C.4: $E(\boldsymbol{R}^{\max})$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.5304	0.5259	0.5222	0.5196	0.5170	0.5124	0.5089	0.5058	0.5042
AcOH-AcOH	0.7481	0.7351	0.7244	0.7154	0.7080	0.6923	0.6793	0.6703	0.6660
CCl_3OH-H_2O	0.7283	0.7185	0.7104	0.7037	0.6891	0.6865	0.6770	0.6702	0.6658
CF_3OH-H_2O	0.7233	0.7145	0.7072	0.7012	0.6963	0.6861	0.6780	0.6719	0.6678
HBr-MeOH	0.2472	0.2480	0.2485	0.2487	0.2489	0.2492	0.2497	0.2499	0.2501
HCl-$MeNH_2$	0.3149	0.3124	0.3102	0.3085	0.3072	0.3061	0.3060	0.3059	0.3055
HCl-MeOH	0.3317	0.3316	0.3313	0.3309	0.3306	0.3306	0.3298	0.3294	0.3293
HF-$MeNH_2$	1.0575	1.0386	1.0211	1.0075	0.9960	0.9708	0.9488	0.9336	0.9246
HF-MeOH	1.0321	1.0180	1.0064	0.9968	0.9889	0.9729	0.9606	0.9514	0.9452
$MeNH_2$-$MeNH_2$	0.5163	0.5139	0.5121	0.5106	0.5095	0.5075	0.5056	0.5042	0.5035
$MeNH_2$-MeOH	0.5156	0.5136	0.5121	0.5110	0.5101	0.5085	0.5070	0.5059	0.5051
$MeNH_2$-Peptide	0.5161	0.5140	0.5123	0.5110	0.5100	0.5077	0.5058	0.5048	0.5042
$MeNH_2$-Pyridine	0.5138	0.5119	0.5104	0.5094	0.5086	0.5074	0.5064	0.5055	0.5049
MeOH-CH_3Cl	0.6861	0.6817	0.6786	0.6763	0.6746	0.6718	0.6693	0.6681	0.6677
MeOH-CH_3F	0.6916	0.6870	0.6834	0.6807	0.6787	0.6749	0.6717	0.6696	0.6683
MeOH-$MeNH_2$	0.7141	0.7050	0.6975	0.6911	0.6859	0.6749	0.6658	0.6589	0.6550
MeOH-MeOH	0.7051	0.6975	0.6914	0.6864	0.6824	0.6745	0.6683	0.6636	0.6609
MeOH-Peptide	0.7129	0.7042	0.6970	0.6911	0.6862	0.6760	0.6667	0.6591	0.6563
MeOH-Pyridine	0.7132	0.7040	0.6964	0.6901	0.6849	0.6743	0.6660	0.6598	0.6561
MeOH-H_2O	0.7033	0.6965	0.6910	0.6866	0.6831	0.6762	0.6706	0.6663	0.6635
Peptide-$MeNH_2$	0.5304	0.5277	0.5254	0.5234	0.5218	0.5188	0.5166	0.5146	0.5133
Peptide-MeOH	0.5319	0.5296	0.5277	0.5263	0.5252	0.5232	0.5217	0.5202	0.5192
Peptide-H_2O	0.5320	0.5302	0.5286	0.5273	0.5263	0.5245	0.5227	0.5214	0.5201
H_2O-$MeNH_2$	0.6991	0.6910	0.6842	0.6787	0.6741	0.6646	0.6570	0.6511	0.6476
H_2O-MeOH	0.6936	0.6866	0.6809	0.6764	0.6728	0.6656	0.6602	0.6560	0.6536
H_2O-Peptide	0.6994	0.6914	0.6849	0.6795	0.6751	0.6662	0.6589	0.6525	0.6640
H_2O-Pyridine	0.6992	0.6909	0.6841	0.6785	0.6739	0.6647	0.6577	0.6524	0.6640
H_2O-H_2O	0.6921	0.6857	0.6807	0.6766	0.6735	0.6672	0.6623	0.6583	0.6559

Table C.5: $\langle \boldsymbol{R}_{yz} \rangle$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.0495	0.0610	0.0706	0.0784	0.0853	0.0996	0.1123	0.1235	0.1292
AcOH-AcOH	-0.0873	-0.0756	-0.0659	-0.0578	-0.0509	-0.0361	-0.0233	-0.0142	-0.0098
CCl_3OH-H_2O	-0.0855	-0.0752	-0.0668	-0.0598	-0.0441	-0.0412	-0.0306	-0.0226	-0.0171
CF_3OH-H_2O	-0.0660	-0.0570	-0.0497	-0.0435	-0.0384	-0.0275	-0.0184	-0.0113	-0.0064
HBr-MeOH	0.2539	0.2706	0.2841	0.2953	0.3045	0.3241	0.3408	0.3544	0.3643
HCl-$MeNH_2$	0.0462	0.0735	0.0966	0.1160	0.1322	0.1666	0.1987	0.2242	0.2407
HCl-MeOH	0.1682	0.1822	0.1936	0.2003	0.2109	0.2109	0.2417	0.2531	0.2613
HF-$MeNH_2$	-0.1553	-0.1460	-0.1384	-0.1319	-0.1264	-0.1140	-0.1028	-0.0946	-0.0892
HF-MeOH	-0.1254	-0.1185	-0.1126	-0.1076	-0.1035	-0.0949	-0.0884	-0.0834	-0.0795
$MeNH_2$-$MeNH_2$	0.1392	0.1462	0.1521	0.1569	0.1609	0.1691	0.1775	0.1827	0.1854
$MeNH_2$-MeOH	0.1438	0.1499	0.1548	0.1587	0.1618	0.1681	0.1738	0.1780	0.1808
$MeNH_2$-Peptide	0.1388	0.1450	0.1501	0.1543	0.1578	0.1676	0.1750	0.1787	0.1810
$MeNH_2$-Pyridine	0.1502	0.1553	0.1595	0.1628	0.1655	0.1707	0.1750	0.1785	0.1808
MeOH-CH_3Cl	0.0184	0.0223	0.0253	0.0275	0.0291	0.0319	0.0345	0.0359	0.0365
MeOH-CH_3F	0.0102	0.0151	0.0190	0.0220	0.0244	0.0287	0.0324	0.0349	0.0365
MeOH-$MeNH_2$	-0.0278	-0.0194	-0.0123	-0.0063	-0.0012	0.0098	0.0195	0.0274	0.0323
MeOH-MeOH	0.0043	0.0117	0.0178	0.0229	0.0271	0.0355	0.0424	0.0478	0.0511
MeOH-Peptide	-0.0091	-0.0008	0.0061	0.0120	0.0170	0.0275	0.0375	0.0461	0.0495
MeOH-Pyridine	-0.0273	-0.0184	-0.0109	-0.0027	0.0019	0.0121	0.0216	0.0283	0.0329
MeOH-H_2O	0.0044	0.0111	0.0166	0.0211	0.0247	0.0321	0.0384	0.0435	0.0469
Peptide-$MeNH_2$	0.1111	0.1198	0.1271	0.1332	0.1383	0.1493	0.1591	0.1671	0.1718
Peptide-MeOH	0.1294	0.1364	0.1422	0.1469	0.1507	0.1586	0.1653	0.1711	0.1749
Peptide-H_2O	0.1085	0.1138	0.1192	0.1235	0.1271	0.1343	0.1419	0.1462	0.1502
H_2O-$MeNH_2$	-0.0211	-0.0134	-0.0068	-0.0013	0.0034	0.0134	0.0221	0.0291	0.0335
H_2O-MeOH	-0.0055	0.0018	0.0078	0.0127	0.0168	0.0249	0.0315	0.0365	0.0397
H_2O-Peptide	-0.0223	-0.0143	-0.0075	-0.0018	0.0030	0.0132	0.0219	0.0297	0.0388
H_2O-Pyridine	-0.0213	-0.0130	-0.0060	-0.0002	0.0047	0.0148	0.0230	0.0293	0.0387
H_2O-H_2O	-0.0047	0.0018	0.0071	0.0114	0.0149	0.0219	0.0278	0.0326	0.0359

Table C.6: $\delta_{\boldsymbol{R}}$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$
$AcNH_2$-$AcNH_2$	0.0660	0.0587	0.0534	0.0494	0.0460	0.0378	0.0250	0.0064
AcOH-AcOH	0.1041	0.0915	0.0806	0.0717	0.0647	0.0496	0.0321	0.0116
CCl_3OH-H_2O	0.0839	0.0762	0.0700	0.0651	0.0533	0.0505	0.0338	0.0126
CF_3OH-H_2O	0.0698	0.0626	0.0577	0.0542	0.0513	0.0432	0.0289	0.0102
HBr-MeOH	0.0158	0.0133	0.0116	0.0103	0.0093	0.0072	0.0045	0.0015
HCl-$MeNH_2$	0.1297	0.1177	0.1073	0.0979	0.0891	0.0672	0.0459	0.0183
HCl-MeOH	0.0165	0.0140	0.0122	0.0109	0.0100	0.0100	0.0047	0.0016
HF-$MeNH_2$	0.0921	0.0851	0.0790	0.0739	0.0698	0.0603	0.0432	0.0158
HF-MeOH	0.0190	0.0165	0.0147	0.0134	0.0123	0.0097	0.0061	0.0023
$MeNH_2$-$MeNH_2$	0.0286	0.0274	0.0264	0.0252	0.0239	0.0189	0.0087	0.0011
$MeNH_2$-MeOH	0.0256	0.0255	0.0250	0.0241	0.0228	0.0175	0.0094	0.0012
$MeNH_2$-Peptide	0.0248	0.0213	0.0185	0.0163	0.0144	0.0076	0.0021	0.0004
$MeNH_2$-Pyridine	0.0305	0.0247	0.0205	0.0172	0.0147	0.0095	0.0044	0.0009
MeOH-CH_3Cl	0.0197	0.0162	0.0138	0.0119	0.0102	0.0063	0.0029	0.0010
MeOH-CH_3F	0.0358	0.0336	0.0314	0.0290	0.0265	0.0189	0.0101	0.0022
MeOH-$MeNH_2$	0.0537	0.0502	0.0478	0.0459	0.0441	0.0379	0.0253	0.0077
MeOH-MeOH	0.0415	0.0386	0.0365	0.0347	0.0329	0.0270	0.0166	0.0046
MeOH-Peptide	0.0476	0.0422	0.0384	0.0355	0.0330	0.0265	0.0151	0.0026
MeOH-Pyridine	0.0522	0.0475	0.0439	0.0409	0.0382	0.0307	0.0196	0.0066
MeOH-H_2O	0.0374	0.0360	0.0347	0.0334	0.0320	0.0263	0.0159	0.0039
Peptide-$MeNH_2$	0.0453	0.0432	0.0420	0.0410	0.0399	0.0351	0.0240	0.0073
Peptide-MeOH	0.0083	0.0074	0.0067	0.0062	0.0057	0.0046	0.0028	0.0010
Peptide-H_2O	0.0223	0.0216	0.0208	0.0200	0.0189	0.0151	0.0091	0.0026
H_2O-$MeNH_2$	0.0486	0.0460	0.0442	0.0427	0.0412	0.0355	0.0233	0.0065
H_2O-MeOH	0.0423	0.0396	0.0375	0.0356	0.0337	0.0273	0.0162	0.0042
H_2O-Peptide	0.0516	0.0469	0.0433	0.0405	0.0380	0.0307	0.0188	0.0060
H_2O-Pyridine	0.0474	0.0433	0.0402	0.0376	0.0352	0.0283	0.0179	0.0058
H_2O-H_2O	0.0377	0.0365	0.0354	0.0341	0.0326	0.0266	0.0156	0.0036

Table C.7: $\langle \boldsymbol{R}^0_{\boldsymbol{yz}} \rangle$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.8802	0.8765	0.8742	0.8728	0.8722	0.8721	0.8747	0.8788	0.8806
AcOH-AcOH	0.8898	0.8843	0.8803	0.8776	0.8759	0.8746	0.8762	0.8810	0.8840
CCl_3OH-H_2O	0.8858	0.8809	0.8775	0.8753	0.8735	0.8737	0.8766	0.8823	0.8857
CF_3OH-H_2O	0.8848	0.8809	0.8783	0.8767	0.8759	0.8761	0.8785	0.8830	0.8858
HBr-MeOH	0.3707	0.3693	0.3683	0.3678	0.3674	0.3668	0.3663	0.3662	0.3660
HCl-$MeNH_2$	0.7721	0.7645	0.7585	0.7538	0.7503	0.7459	0.7472	0.7514	0.7560
HCl-MeOH	0.3719	0.3703	0.3692	0.3685	0.3680	0.3680	0.3667	0.3664	0.3662
HF-$MeNH_2$	0.7657	0.7593	0.7543	0.7505	0.7479	0.7451	0.7465	0.7512	0.7552
HF-MeOH	0.3747	0.3728	0.3714	0.3705	0.3698	0.3686	0.3676	0.367	0.3664
$MeNH_2$-$MeNH_2$	0.7151	0.7145	0.7144	0.7145	0.7147	0.7155	0.7172	0.7188	0.7190
$MeNH_2$-MeOH	0.8787	0.8783	0.8782	0.8784	0.8786	0.8796	0.8814	0.8838	0.8840
$MeNH_2$-Peptide	0.7170	0.7158	0.7149	0.7143	0.7139	0.7129	0.7123	0.7123	0.7123
$MeNH_2$-Pyridine	0.7214	0.7207	0.7203	0.7200	0.7199	0.7198	0.7200	0.7204	0.7204
MeOH-CH_3Cl	0.6324	0.6309	0.6300	0.6296	0.6294	0.6298	0.6307	0.6313	0.6316
MeOH-CH_3F	1.0503	1.0502	1.0504	1.0508	1.0513	1.0530	1.0555	1.0584	1.0589
MeOH-$MeNH_2$	0.7516	0.7485	0.7466	0.7457	0.7455	0.7464	0.7487	0.7524	0.7545
MeOH-MeOH	0.8493	0.8474	0.8466	0.8463	0.8464	0.8470	0.8486	0.8512	0.8524
MeOH-Peptide	0.8167	0.8120	0.8091	0.8074	0.8064	0.8051	0.8047	0.8058	0.8064
MeOH-Pyridine	0.7369	0.7329	0.7299	0.7279	0.7264	0.7242	0.7239	0.7255	0.7266
MeOH-H_2O	0.8593	0.8582	0.8580	0.8583	0.8587	0.8603	0.8626	0.8658	0.8668
Peptide-$MeNH_2$	0.7428	0.7405	0.7394	0.7390	0.7390	0.7401	0.7427	0.7470	0.7486
Peptide-MeOH	0.3202	0.3197	0.3193	0.3191	0.3189	0.3186	0.3183	0.3181	0.3179
Peptide-H_2O	0.8764	0.8760	0.8760	0.8762	0.8765	0.8773	0.8785	0.8801	0.8804
H_2O-$MeNH_2$	0.7506	0.7478	0.7462	0.7455	0.7453	0.7460	0.7482	0.7518	0.7535
H_2O-MeOH	0.8703	0.8687	0.8680	0.8679	0.8680	0.8690	0.8711	0.8741	0.8754
H_2O-Peptide	0.8825	0.8797	0.8781	0.8772	0.8768	0.8771	0.8790	0.8825	0.8840
H_2O-Pyridine	0.7355	0.7319	0.7293	0.7274	0.7262	0.7243	0.7242	0.7258	0.7267
H_2O-H_2O	0.8785	0.8773	0.8768	0.8769	0.8772	0.8788	0.8814	0.8848	0.8858

Table C.8: $R_z^{\max}$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.4231	0.4247	0.4259	0.4269	0.4274	0.4290	0.4305	0.4319	0.4325
AcOH-AcOH	0.4212	0.4241	0.4263	0.4279	0.4292	0.4319	0.4343	0.4361	0.4372
CCl_3OH-H_2O	0.4216	0.4238	0.4255	0.4267	0.4291	0.4296	0.4314	0.4334	0.4341
CF_3OH-H_2O	0.4443	0.4467	0.4484	0.4497	0.4505	0.4522	0.4540	0.4557	0.4568
HBr-MeOH	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641
HCl-$MeNH_2$	0.3039	0.3098	0.3142	0.3177	0.3205	0.3259	0.3301	0.3333	0.3354
HCl-MeOH	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641
HF-$MeNH_2$	0.3089	0.3131	0.3165	0.3191	0.3212	0.3256	0.3296	0.3329	0.3350
HF-MeOH	0.8641	0.8641	0.9695	0.8641	0.8641	0.8641	0.8641	0.8641	0.8641
$MeNH_2$-$MeNH_2$	0.3875	0.3884	0.3891	0.3897	0.3902	0.3913	0.3925	0.3932	0.3933
$MeNH_2$-MeOH	0.4290	0.4294	0.4296	0.4299	0.4301	0.4306	0.4311	0.4315	0.4316
$MeNH_2$-Peptide	0.5708	0.5727	0.5741	0.5752	0.5760	0.5783	0.5797	0.5800	0.5802
$MeNH_2$-Pyridine	0.6223	0.6263	0.6288	0.6310	0.6325	0.6350	0.6365	0.6370	0.6372
MeOH-CH_3Cl	0.0553	0.0582	0.0604	0.0619	0.0631	0.0653	0.0669	0.0678	0.0681
MeOH-CH_3F	0.5020	0.5024	0.5027	0.5029	0.5032	0.5036	0.5041	0.5044	0.5046
MeOH-$MeNH_2$	0.3233	0.3254	0.3270	0.3283	0.3293	0.3315	0.3335	0.3353	0.3362
MeOH-MeOH	0.4280	0.4290	0.4297	0.4303	0.4307	0.4318	0.4328	0.4336	0.4341
MeOH-Peptide	0.4257	0.4269	0.4279	0.4287	0.4294	0.4307	0.4322	0.4333	0.4337
MeOH-Pyridine	0.3353	0.3374	0.3391	0.3404	0.3414	0.3435	0.3456	0.3472	0.3482
MeOH-H_2O	0.4397	0.4405	0.4411	0.4415	0.4419	0.4428	0.4438	0.4446	0.4450
Peptide-$MeNH_2$	0.3306	0.3321	0.3333	0.3342	0.3350	0.3367	0.3384	0.3401	0.3411
Peptide-MeOH	0.8802	0.8802	0.8802	0.8801	0.8801	0.9921	0.8788	0.8776	0.8766
Peptide-H_2O	0.6176	0.6182	0.6186	0.6190	0.6194	0.6204	0.6215	0.6222	0.6227
H_2O-$MeNH_2$	0.3331	0.3350	0.3365	0.3376	0.3386	0.3406	0.3425	0.3442	0.3451
H_2O-MeOH	0.4262	0.4269	0.4276	0.4281	0.4285	0.4295	0.4305	0.4313	0.4317
H_2O-Peptide	0.4265	0.4275	0.4285	0.4292	0.4297	0.4310	0.4322	0.4333	0.4339
H_2O-Pyridine	0.3394	0.3413	0.3427	0.3439	0.3448	0.3469	0.3488	0.3504	0.3513
H_2O-H_2O	0.4333	0.4340	0.4346	0.4350	0.4354	0.4363	0.4372	0.4379	0.4383

Table C.9: $E(\boldsymbol{R}^{\max})$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.8134	0.8132	0.8134	0.8141	0.8151	0.8194	0.8283	0.8419	0.8474
AcOH-AcOH	0.8046	0.8046	0.8056	0.8072	0.8089	0.8149	0.8252	0.8395	0.8484
CCl_3OH-H_2O	0.8208	0.8200	0.8204	0.8217	0.8290	0.8313	0.8455	0.8626	0.8749
CF_3OH-H_2O	0.8293	0.8286	0.8288	0.8298	0.8314	0.8381	0.8502	0.8650	0.8746
HBr-MeOH	0.2716	0.2710	0.2706	0.2703	0.2700	0.2696	0.2694	0.2694	0.2706
HCl-$MeNH_2$	0.4862	0.4822	0.4800	0.4791	0.4794	0.4837	0.4924	0.5054	0.5150
HCl-MeOH	0.2724	0.2718	0.2712	0.2708	0.2705	0.2705	0.2697	0.2696	0.2694
HF-$MeNH_2$	0.4989	0.4937	0.4904	0.4883	0.4870	0.4872	0.4944	0.5076	0.5163
HF-MeOH	0.2752	0.2741	0.2688	0.2725	0.272	0.2711	0.2706	0.2703	0.2698
$MeNH_2$-$MeNH_2$	0.4783	0.4779	0.4779	0.4782	0.4787	0.4808	0.4854	0.4892	0.4898
$MeNH_2$-MeOH	0.8607	0.8607	0.8611	0.8619	0.8629	0.8668	0.8727	0.8792	0.8804
$MeNH_2$-Peptide	0.6599	0.6590	0.6585	0.6583	0.6584	0.6602	0.6623	0.6633	0.6635
$MeNH_2$-Pyridine	0.5367	0.5368	0.5370	0.5374	0.5378	0.5394	0.5415	0.5435	0.5439
MeOH-CH_3Cl	0.2924	0.2919	0.2917	0.2917	0.2919	0.2928	0.2937	0.2944	0.2947
MeOH-CH_3F	1.3629	1.3643	1.3663	1.3686	1.3711	1.3790	1.3884	1.3975	1.4002
MeOH-$MeNH_2$	0.4985	0.4964	0.4953	0.4950	0.4952	0.4979	0.5040	0.5125	0.5169
MeOH-MeOH	0.7976	0.7969	0.7972	0.7979	0.7990	0.8031	0.8102	0.8189	0.8229
MeOH-Peptide	0.7543	0.7512	0.7496	0.7489	0.7488	0.7505	0.7560	0.7634	0.7654
MeOH-Pyridine	0.5345	0.5329	0.5320	0.5317	0.5319	0.5342	0.5394	0.5463	0.5502
MeOH-H_2O	0.8120	0.8122	0.8132	0.8146	0.8162	0.8214	0.8294	0.8386	0.8422
Peptide-$MeNH_2$	0.4983	0.4967	0.4958	0.4957	0.4959	0.4984	0.5041	0.5126	0.5168
Peptide-MeOH	0.2618	0.2614	0.2611	0.2609	0.2608	0.2640	0.2603	0.2601	0.2597
Peptide-H_2O	0.8506	0.8510	0.8517	0.8526	0.8536	0.8564	0.8604	0.8652	0.8675
H_2O-$MeNH_2$	0.4994	0.4975	0.4964	0.4961	0.4964	0.4989	0.5047	0.5130	0.5168
H_2O-MeOH	0.8317	0.8313	0.8317	0.8326	0.8339	0.8386	0.8467	0.8559	0.8598
H_2O-Peptide	0.8240	0.8231	0.8229	0.8232	0.8240	0.8278	0.8357	0.8457	0.8499
H_2O-Pyridine	0.5358	0.5343	0.5335	0.5333	0.5335	0.5357	0.5406	0.5471	0.5505
H_2O-H_2O	0.8437	0.8437	0.8445	0.8459	0.8475	0.8530	0.8617	0.8714	0.8750

Table C.10: $\langle \boldsymbol{R}_{yz} \rangle$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.4005	0.4088	0.4152	0.4202	0.4243	0.4333	0.4448	0.4606	0.4655
AcOH-AcOH	0.3745	0.3856	0.3958	0.4046	0.4117	0.4269	0.4437	0.4606	0.4704
CCl_3OH-H_2O	0.3650	0.3738	0.3814	0.3874	0.4001	0.4026	0.4167	0.4349	0.4422
CF_3OH-H_2O	0.4019	0.4098	0.4158	0.4202	0.4235	0.4312	0.4437	0.4588	0.4661
HBr-MeOH	0.7337	0.7445	0.7525	0.7584	0.7628	0.7720	0.7823	0.7910	0.7952
HCl-$MeNH_2$	0.1942	0.2073	0.2210	0.2350	0.2487	0.2829	0.3150	0.3535	0.3746
HCl-MeOH	0.7329	0.7431	0.7508	0.7567	0.7612	0.7612	0.7805	0.7899	0.7943
HF-$MeNH_2$	0.2613	0.2678	0.2748	0.2813	0.2871	0.3004	0.3211	0.3539	0.3718
HF-MeOH	0.7386	0.745	0.7501	0.7542	0.7576	0.7656	0.7761	0.7874	0.7938
$MeNH_2$-$MeNH_2$	0.3906	0.3934	0.3957	0.3977	0.3998	0.4065	0.4196	0.4291	0.4301
$MeNH_2$-MeOH	0.4105	0.4117	0.4128	0.4141	0.4155	0.4204	0.4271	0.4333	0.4341
$MeNH_2$-Peptide	0.5878	0.5914	0.5943	0.5967	0.5989	0.6072	0.6142	0.6165	0.6168
$MeNH_2$-Pyridine	0.5702	0.5761	0.5808	0.5846	0.5877	0.5943	0.6009	0.6055	0.6064
MeOH-CH_3Cl	0.1991	0.2053	0.2101	0.2140	0.2172	0.2237	0.2275	0.2291	0.2297
MeOH-CH_3F	0.5010	0.5032	0.5051	0.5069	0.5087	0.5137	0.5188	0.5228	0.5242
MeOH-$MeNH_2$	0.3021	0.3080	0.3126	0.3161	0.3191	0.3274	0.3424	0.3632	0.3719
MeOH-MeOH	0.4257	0.4298	0.4327	0.4351	0.4372	0.4434	0.4533	0.4639	0.4677
MeOH-Peptide	0.4102	0.4165	0.4213	0.4252	0.4284	0.4362	0.4489	0.4618	0.4641
MeOH-Pyridine	0.2234	0.2295	0.2347	0.2392	0.2431	0.2533	0.2671	0.2824	0.2896
MeOH-H_2O	0.4334	0.4360	0.4377	0.4390	0.4403	0.4451	0.4538	0.4635	0.4664
Peptide-$MeNH_2$	0.3073	0.3115	0.3144	0.3165	0.3184	0.3243	0.3370	0.3556	0.3641
Peptide-MeOH	0.8269	0.8303	0.8330	0.8353	0.8372	0.8421	0.8483	0.8549	0.8578
Peptide-H_2O	0.6145	0.6156	0.6165	0.6173	0.6183	0.6217	0.6270	0.6324	0.6344
H_2O-$MeNH_2$	0.3158	0.3210	0.3249	0.3279	0.3305	0.3383	0.3529	0.3727	0.3799
H_2O-MeOH	0.4069	0.4105	0.4133	0.4155	0.4176	0.4236	0.4334	0.4431	0.4462
H_2O-Peptide	0.4118	0.4175	0.4219	0.4255	0.4286	0.4363	0.4474	0.4590	0.4630
H_2O-Pyridine	0.2321	0.2377	0.2424	0.2464	0.2499	0.2594	0.2723	0.2867	0.2929
H_2O-H_2O	0.4141	0.4166	0.4183	0.4198	0.4212	0.4262	0.4350	0.4441	0.4466

Table C.11: $\delta_{\boldsymbol{u}}$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$
$AcNH_2$-$AcNH_2$	0.0391	0.0336	0.0287	0.0244	0.0208	0.0132	0.0073	0.0024
AcOH-AcOH	0.0630	0.0545	0.0471	0.0405	0.0348	0.0219	0.0109	0.0035
CCl_3OH-H_2O	0.0496	0.0427	0.0367	0.0314	0.0192	0.0169	0.0088	0.0033
CF_3OH-H_2O	0.0446	0.0381	0.0325	0.0276	0.0234	0.0147	0.0079	0.0029
HBr-MeOH	0.0099	0.0091	0.0079	0.0068	0.0057	0.0032	0.0017	0.0006
HCl-$MeNH_2$	0.0290	0.0258	0.0231	0.0208	0.0188	0.0140	0.0078	0.0033
HCl-MeOH	0.0205	0.0180	0.0155	0.0132	0.0111	0.0111	0.0038	0.0014
HF-$MeNH_2$	0.0780	0.0676	0.0588	0.0514	0.0448	0.0292	0.0147	0.0052
HF-MeOH	0.0532	0.0451	0.0381	0.0321	0.0270	0.0166	0.0090	0.0034
$MeNH_2$-$MeNH_2$	0.0259	0.0211	0.0173	0.0142	0.0118	0.0074	0.0038	0.0012
$MeNH_2$-MeOH	0.0205	0.0166	0.0135	0.0112	0.0094	0.0062	0.0034	0.0013
$MeNH_2$-Peptide	0.0233	0.0190	0.0156	0.0129	0.0108	0.0060	0.0026	0.0010
$MeNH_2$-Pyr	0.0200	0.0157	0.0122	0.0096	0.0077	0.0047	0.0027	0.0010
MeOH-CH_3Cl	0.0206	0.0155	0.0117	0.0088	0.0068	0.0038	0.0013	0.0003
MeOH-CH_3F	0.0223	0.0177	0.0141	0.0114	0.0094	0.0058	0.0029	0.0010
MeOH-$MeNH_2$	0.0525	0.0452	0.0388	0.0332	0.0284	0.0180	0.0095	0.0033
MeOH-MeOH	0.0403	0.0336	0.0280	0.0234	0.0196	0.0121	0.0064	0.0021
MeOH-Peptide	0.0497	0.0425	0.0363	0.0310	0.0265	0.0172	0.0090	0.0023
MeOH-Pyr	0.0507	0.0434	0.0369	0.0340	0.0283	0.0167	0.0093	0.0032
MeOH-H_2O	0.0370	0.0307	0.0255	0.0213	0.0180	0.0113	0.0062	0.0022
Peptide-$MeNH_2$	0.0343	0.0292	0.0248	0.0209	0.0177	0.0110	0.0061	0.0022
Peptide-MeOH	0.0256	0.0210	0.0173	0.0143	0.0119	0.0076	0.0046	0.0018
Peptide-H_2O	0.0244	0.0200	0.0164	0.0137	0.0115	0.0076	0.0046	0.0019
H_2O-$MeNH_2$	0.0463	0.0395	0.0337	0.0286	0.0243	0.0152	0.0082	0.0029
H_2O-MeOH	0.0365	0.0302	0.0250	0.0208	0.0173	0.0106	0.0057	0.0019
H_2O-Peptide	0.0438	0.0371	0.0314	0.0265	0.0224	0.0142	0.0079	0.0023
H_2O-Pyr	0.0450	0.0382	0.0323	0.0273	0.0230	0.0142	0.0076	0.0028
H_2O-H_2O	0.0337	0.0278	0.0230	0.0191	0.0160	0.0100	0.0055	0.0019

Table C.12: $\langle u^{-1} \rangle$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.7817	0.7794	0.7774	0.7757	0.7742	0.7711	0.7687	0.7666	0.7655
AcOH-AcOH	0.8634	0.8592	0.8557	0.8525	0.8498	0.8437	0.8385	0.8350	0.8333
CCl_3OH-H_2O	0.8536	0.8506	0.8479	0.8455	0.8400	0.8390	0.8355	0.8329	0.8313
CF_3OH-H_2O	0.8541	0.8511	0.8485	0.8462	0.8443	0.8403	0.8372	0.8348	0.8333
HBr-MeOH	0.6026	0.6027	0.6026	0.6024	0.6022	0.6017	0.6015	0.6013	0.6012
HCl-$MeNH_2$	0.6513	0.6508	0.6504	0.6499	0.6496	0.6486	0.6470	0.6458	0.6447
HCl-MeOH	0.6671	0.6665	0.6658	0.6651	0.6645	0.6645	0.6623	0.6616	0.6611
HF-$MeNH_2$	0.9491	0.9437	0.9393	0.9355	0.9321	0.9243	0.9173	0.9126	0.9098
HF-MeOH	0.9457	0.9416	0.9380	0.9349	0.9323	0.9270	0.9231	0.9202	0.9183
$MeNH_2$-$MeNH_2$	0.7651	0.7634	0.7619	0.7608	0.7599	0.7583	0.7569	0.7558	0.7553
$MeNH_2$-MeOH	0.7644	0.7629	0.7616	0.7608	0.7601	0.7588	0.7577	0.7568	0.7562
$MeNH_2$-Peptide	0.7647	0.7631	0.7619	0.7609	0.7601	0.7583	0.7569	0.7562	0.7557
$MeNH_2$-Pyr	0.7625	0.7612	0.7602	0.7594	0.7588	0.7579	0.7572	0.7565	0.7561
MeOH-CH_3Cl	0.8425	0.8406	0.8392	0.8381	0.8374	0.8363	0.8352	0.8348	0.8346
MeOH-CH_3F	0.8452	0.8431	0.8415	0.8403	0.8393	0.8377	0.8363	0.8354	0.8349
MeOH-$MeNH_2$	0.8525	0.8494	0.8464	0.8440	0.8418	0.8373	0.8336	0.8307	0.8291
MeOH-MeOH	0.8507	0.8477	0.8452	0.8431	0.8413	0.8379	0.8354	0.8333	0.8323
MeOH-Peptide	0.8530	0.8497	0.8469	0.8445	0.8425	0.8383	0.8345	0.8313	0.8302
MeOH-Pyr	0.8523	0.8490	0.8462	0.8439	0.8417	0.8371	0.8338	0.8311	0.8296
MeOH-H_2O	0.8498	0.8470	0.8447	0.8428	0.8413	0.8383	0.8360	0.8341	0.8330
Peptide-$MeNH_2$	0.7814	0.7795	0.7779	0.7764	0.7752	0.7728	0.7710	0.7693	0.7684
Peptide-MeOH	0.7812	0.7795	0.7781	0.7770	0.7761	0.7744	0.7732	0.7720	0.7712
Peptide-H_2O	0.7810	0.7793	0.7780	0.7770	0.7761	0.7746	0.7734	0.7722	0.7714
H_2O-$MeNH_2$	0.8458	0.8427	0.8401	0.8378	0.8359	0.8330	0.8288	0.8264	0.8249
H_2O-MeOH	0.8444	0.8415	0.8392	0.8372	0.8356	0.8326	0.8303	0.8285	0.8275
H_2O-Peptide	0.8460	0.8429	0.8403	0.8381	0.8362	0.8325	0.8296	0.8269	0.8258
H_2O-Pyr	0.8459	0.8428	0.8402	0.8379	0.8360	0.8321	0.8291	0.8269	0.8255
H_2O-H_2O	0.8440	0.8413	0.8391	0.8374	0.8360	0.8332	0.8312	0.8294	0.8284

Table C.13: u_{max} for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	1.3534	1.3573	1.3605	1.3631	1.3653	1.3696	1.3732	1.3765	1.3783
AcOH-AcOH	1.2172	1.2240	1.2295	1.2341	1.2379	1.2459	1.2524	1.2573	1.2597
CCl_3OH-H_2O	1.2233	1.2286	1.2329	1.2364	1.2440	1.2454	1.2502	1.2540	1.2565
CF_3OH-H_2O	1.2262	1.2309	1.2346	1.2377	1.2403	1.2454	1.2495	1.2528	1.2551
HBr-MeOH	1.7167	1.7173	1.7179	1.7185	1.7189	1.7196	1.7197	1.7203	1.7207
HCl-$MeNH_2$	1.5894	1.5944	1.5981	1.6009	1.6027	1.6060	1.6096	1.6127	1.6154
HCl-MeOH	1.5601	1.5617	1.5632	1.5645	1.5656	1.5656	1.5690	1.5707	1.5719
HF-$MeNH_2$	1.1023	1.1101	1.1164	1.1217	1.1261	1.1358	1.1444	1.1504	1.1540
HF-MeOH	1.1107	1.1157	1.1199	1.1234	1.1263	1.1320	1.1365	1.1399	1.1423
$MeNH_2$-$MeNH_2$	1.3692	1.3715	1.3734	1.3748	1.3759	1.3781	1.3806	1.3821	1.3829
$MeNH_2$-MeOH	1.3703	1.3722	1.3736	1.3747	1.3755	1.3772	1.3789	1.3801	1.3810
$MeNH_2$-Peptide	1.3694	1.3715	1.3730	1.3742	1.3752	1.3779	1.3802	1.3813	1.3820
$MeNH_2$-Pyr	1.3719	1.3738	1.3751	1.3761	1.3769	1.3783	1.3796	1.3806	1.3813
MeOH-CH_3Cl	1.2435	1.2456	1.2471	1.2482	1.2490	1.2504	1.2518	1.2524	1.2526
MeOH-CH_3F	1.2408	1.2431	1.2448	1.2460	1.2469	1.2488	1.2504	1.2514	1.2522
MeOH-$MeNH_2$	1.2308	1.2355	1.2394	1.2427	1.2454	1.2512	1.2558	1.2596	1.2618
MeOH-MeOH	1.2340	1.2379	1.2410	1.2435	1.2454	1.2493	1.2525	1.2549	1.2564
MeOH-Peptide	1.2307	1.2352	1.2389	1.2419	1.2443	1.2494	1.2543	1.2583	1.2599
MeOH-Pyr	1.2312	1.2360	1.2399	1.2432	1.2458	1.2513	1.2556	1.2589	1.2610
MeOH-H_2O	1.2351	1.2385	1.2412	1.2434	1.2451	1.2484	1.2513	1.2536	1.2551
Peptide-$MeNH_2$	1.3549	1.3578	1.3603	1.3623	1.3639	1.3672	1.3700	1.3723	1.3738
Peptide-MeOH	1.3553	1.3575	1.3593	1.3606	1.3616	1.3637	1.3655	1.3671	1.3683
Peptide-H_2O	1.3550	1.3570	1.3586	1.3597	1.3607	1.3625	1.3643	1.3658	1.3671
H_2O-$MeNH_2$	1.2385	1.2428	1.2463	1.2493	1.2517	1.2566	1.2606	1.2639	1.2659
H_2O-MeOH	1.2407	1.2443	1.2472	1.2495	1.2513	1.2549	1.2578	1.2599	1.2614
H_2O-Peptide	1.2380	1.2422	1.2457	1.2485	1.2507	1.2553	1.2592	1.2628	1.2645
H_2O-Pyr	1.2383	1.2427	1.2463	1.2492	1.2516	1.2564	1.2601	1.2630	1.2650
H_2O-H_2O	1.2413	1.2446	1.2471	1.2491	1.2507	1.2537	1.2564	1.2585	1.2599

Table C.14: $\langle u \rangle$ for the σ_{X-H} LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	1.7297	1.7328	1.7349	1.7361	1.7367	1.7338	1.7411	1.7304	1.7015
AcOH-AcOH	1.7133	1.7202	1.7246	1.7274	1.7290	1.7297	1.7219	1.7286	1.6926
CCl_3OH-H_2O	1.7012	1.7059	1.7080	1.7085	1.7029	1.7004	1.6826	1.6549	1.6364
CF_3OH-H_2O	1.6893	1.6929	1.6946	1.6947	1.6939	1.6875	1.6728	1.6511	1.6365
HBr-MeOH	1.6912	1.6862	1.6822	1.6788	1.6759	1.6694	1.6623	1.6551	1.6519
HCl-$MeNH_2$	2.0303	2.0446	2.0533	2.0573	2.0573	2.0443	2.0181	1.9790	1.9486
HCl-MeOH	1.6844	1.6811	1.6783	1.6759	1.6738	1.6738	1.6618	1.6549	1.6517
HF-$MeNH_2$	1.9532	1.9696	1.9820	1.9912	1.9978	2.0064	1.9994	1.9697	1.9461
HF-MeOH	1.6677	1.6688	1.6694	1.6697	1.6695	1.6676	1.6625	1.6555	1.6517
$MeNH_2$-$MeNH_2$	1.9838	1.9835	1.9827	1.9814	1.9798	1.9740	1.9603	1.9521	1.9455
$MeNH_2$-MeOH	1.6842	1.6836	1.6827	1.6814	1.6799	1.6741	1.6646	1.6527	1.6503
$MeNH_2$-Peptide	1.7098	1.7094	1.7089	1.7083	1.7075	1.7022	1.6958	1.6934	1.6929
$MeNH_2$-Pyr	1.8931	1.8925	1.8918	1.8910	1.8900	1.8861	1.8803	1.8751	1.8737
MeOH-CH_3Cl	2.2416	2.2439	2.2449	2.2448	2.2439	2.2391	2.2336	2.2312	2.2292
MeOH-CH_3F	1.4479	1.4469	1.4456	1.4441	1.4423	1.4362	1.4290	1.4287	1.4178
MeOH-$MeNH_2$	1.9850	1.9912	1.9949	1.9968	1.9973	1.9939	1.9818	1.9580	1.9452
MeOH-MeOH	1.6869	1.6879	1.6879	1.6872	1.6861	1.6813	1.6710	1.6562	1.6498
MeOH-Peptide	1.7167	1.7196	1.7215	1.7224	1.7227	1.7207	1.7107	1.6949	1.6921
MeOH-Pyr	1.8922	1.8983	1.9026	1.9053	1.9070	1.9068	1.8992	1.8833	1.8737
MeOH-H_2O	1.6854	1.6852	1.6839	1.6820	1.6797	1.6724	1.6598	1.6429	1.6369
Peptide-$MeNH_2$	1.9927	1.9961	1.9978	1.9983	1.9979	1.9934	1.9806	1.9697	1.9445
Peptide-MeOH	1.6704	1.6693	1.6681	1.6668	1.6656	1.6620	1.6570	1.6518	1.6502
Peptide-H_2O	1.6670	1.6656	1.6640	1.6623	1.6606	1.6558	1.6496	1.6468	1.6369
H_2O-$MeNH_2$	1.9858	1.9911	1.9943	1.9959	1.9962	1.9928	1.9804	1.9563	1.9455
H_2O-MeOH	1.6861	1.6872	1.6872	1.6865	1.6853	1.6805	1.6699	1.6552	1.6496
H_2O-H_2O	1.6845	1.6844	1.6833	1.6815	1.6794	1.6721	1.6589	1.6414	1.6361

Table C.15: $\delta_{\boldsymbol{u}}$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$
$AcNH_2$-$AcNH_2$	0.0217	0.0244	0.0262	0.0273	0.0278	0.0269	0.0211	0.0070
AcOH-AcOH	0.0242	0.0261	0.0285	0.0299	0.0305	0.0296	0.0229	0.0109
CCl_3OH-H_2O	0.0397	0.0448	0.0475	0.0486	0.0457	0.0440	0.0314	0.0117
CF_3OH-H_2O	0.0343	0.0383	0.0403	0.0410	0.0406	0.0363	0.0253	0.0092
HBr-MeOH	0.0171	0.0157	0.0146	0.0136	0.0127	0.0104	0.0066	0.0022
HCl-$MeNH_2$	0.0421	0.0428	0.0442	0.0464	0.0495	0.0482	0.0376	0.0163
HCl-MeOH	0.0164	0.0154	0.0146	0.0138	0.0130	0.0130	0.0067	0.0022
HF-$MeNH_2$	0.0171	0.0182	0.0238	0.0304	0.0352	0.0409	0.0347	0.0135
HF-MeOH	0.0144	0.0150	0.0152	0.0151	0.0148	0.0127	0.0083	0.0029
$MeNH_2$-$MeNH_2$	0.0197	0.0207	0.0210	0.0209	0.0204	0.0173	0.0085	0.0012
$MeNH_2$-MeOH	0.0203	0.0209	0.0210	0.0207	0.0200	0.0161	0.0091	0.0011
$MeNH_2$-Peptide	0.0107	0.0109	0.0110	0.0107	0.0102	0.0063	0.0017	0.0003
$MeNH_2$-Pyr	0.0116	0.0117	0.0116	0.0113	0.0107	0.0081	0.0041	0.0006
MeOH-CH_3Cl	0.0061	0.0078	0.0087	0.0088	0.0084	0.0057	0.0027	0.0010
MeOH-CH_3F	0.0223	0.0222	0.0215	0.0205	0.0192	0.0142	0.0078	0.0023
MeOH-$MeNH_2$	0.0202	0.0259	0.0295	0.0315	0.0323	0.0308	0.0226	0.0071
MeOH-MeOH	0.0239	0.0255	0.0261	0.0260	0.0255	0.0224	0.0146	0.0041
MeOH-Peptide	0.0192	0.0214	0.0227	0.0234	0.0235	0.0216	0.0134	0.0019
MeOH-Pyr	0.0139	0.0178	0.0211	0.0231	0.0243	0.0238	0.0176	0.0059
MeOH-H_2O	0.0310	0.0317	0.0313	0.0304	0.0292	0.0245	0.0153	0.0036
Peptide-$MeNH_2$	0.0230	0.0272	0.0297	0.0310	0.0316	0.0300	0.0219	0.0076
Peptide-MeOH	0.0120	0.0117	0.0115	0.0108	0.0102	0.0080	0.0047	0.0012
Peptide-H_2O	0.0181	0.0179	0.0174	0.0167	0.0159	0.0130	0.0082	0.0026
H_2O-$MeNH_2$	0.0204	0.0255	0.0287	0.0305	0.0313	0.0298	0.0214	0.0059
H_2O-MeOH	0.0234	0.0250	0.0257	0.0256	0.0251	0.0219	0.0138	0.0035
H_2O-H_2O	0.0308	0.0316	0.0314	0.0307	0.0295	0.0247	0.0150	0.0032

Table C.16: $\langle u^{-1} \rangle$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	0.8397	0.8387	0.8381	0.8379	0.8378	0.8388	0.8427	0.8495	0.8520
AcOH-AcOH	0.8400	0.8381	0.8371	0.8367	0.8367	0.8378	0.8415	0.8485	0.8527
CCl_3OH-H_2O	0.8433	0.8417	0.8410	0.8409	0.8432	0.8442	0.8506	0.8597	0.8655
CF_3OH-H_2O	0.8466	0.8453	0.8448	0.8448	0.8452	0.8478	0.8533	0.8608	0.8654
HBr-MeOH	0.8569	0.8578	0.8585	0.8591	0.8596	0.8601	0.8628	0.8649	0.8660
HCl-$MeNH_2$	0.7132	0.7087	0.7059	0.7043	0.7038	0.7057	0.7110	0.7199	0.7266
HCl-MeOH	0.8573	0.8579	0.8585	0.8590	0.8594	0.8594	0.8627	0.8649	0.8660
HF-$MeNH_2$	0.7237	0.7189	0.7154	0.7130	0.7114	0.7099	0.7131	0.7217	0.7274
HF-MeOH	0.8584	0.8582	0.8582	0.8583	0.8586	0.8597	0.8619	0.8645	0.8659
$MeNH_2$-$MeNH_2$	0.7193	0.7191	0.7190	0.7192	0.7195	0.7210	0.7245	0.7276	0.7279
$MeNH_2$-MeOH	0.8569	0.8568	0.8569	0.8572	0.8576	0.8595	0.8627	0.8664	0.8669
$MeNH_2$-Peptide	0.8486	0.8486	0.8488	0.8490	0.8493	0.8512	0.8533	0.8541	0.8542
$MeNH_2$-Pyr	0.7478	0.7478	0.7480	0.7482	0.7484	0.7496	0.7511	0.7526	0.7528
MeOH-CH_3Cl	0.6200	0.6195	0.6193	0.6193	0.6195	0.6205	0.6215	0.6221	0.6224
MeOH-CH_3F	0.9983	0.9986	0.9992	0.9999	1.0007	1.0035	1.0070	1.0106	1.0114
MeOH-$MeNH_2$	0.7178	0.7160	0.7148	0.7143	0.7142	0.7153	0.7189	0.7249	0.7279
MeOH-MeOH	0.8542	0.8537	0.8537	0.8539	0.8544	0.8562	0.8600	0.8650	0.8670
MeOH-Peptide	0.8440	0.8431	0.8427	0.8426	0.8427	0.8440	0.8481	0.8535	0.8545
MeOH-Pyr	0.7462	0.7445	0.7433	0.7427	0.7424	0.7430	0.7458	0.7504	0.7529
MeOH-H_2O	0.8498	0.8498	0.8501	0.8506	0.8513	0.8538	0.8581	0.8634	0.8653
Peptide-$MeNH_2$	0.7172	0.7158	0.7151	0.7147	0.7147	0.7158	0.7194	0.7256	0.7281
Peptide-MeOH	0.8608	0.8610	0.8613	0.8616	0.8620	0.8631	0.8646	0.8662	0.8668
Peptide-H_2O	0.8564	0.8566	0.8569	0.8574	0.8578	0.8593	0.8615	0.8643	0.8653
H_2O-$MeNH_2$	0.7180	0.7163	0.7153	0.7148	0.7147	0.7157	0.7193	0.7253	0.7278
H_2O-MeOH	0.8545	0.8540	0.8539	0.8542	0.8546	0.8565	0.8604	0.8653	0.8671
H_2O-H_2O	0.8503	0.8502	0.8504	0.8509	0.8516	0.8541	0.8586	0.8640	0.8656

Table C.17: $u_{\max}$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	1.1583	1.1567	1.1553	1.1541	1.1529	1.1496	1.1456	1.1416	1.1397
AcOH-AcOH	1.1691	1.1661	1.1632	1.1607	1.1585	1.1534	1.1478	1.1430	1.1401
CCl_3OH-H_2O	1.1628	1.1611	1.1594	1.1578	1.1533	1.1522	1.1468	1.1429	1.1384
CF_3OH-H_2O	1.1580	1.1569	1.1559	1.1548	1.1537	1.1503	1.1459	1.1419	1.1386
HBr-MeOH	1.1408	1.1402	1.1397	1.1392	1.1388	1.1376	1.1363	1.1353	1.1344
HCl-$MeNH_2$	1.3960	1.3983	1.3986	1.3976	1.3957	1.3893	1.3813	1.3723	1.3661
HCl-MeOH	1.1419	1.1413	1.1407	1.1402	1.1396	1.1396	1.1368	1.1355	1.1344
HF-$MeNH_2$	1.3883	1.3915	1.3925	1.3924	1.3919	1.3884	1.3801	1.3705	1.3642
HF-MeOH	1.1450	1.1442	1.1434	1.1427	1.1420	1.1401	1.1381	1.1362	1.1347
$MeNH_2$-$MeNH_2$	1.3710	1.3713	1.3712	1.3707	1.3701	1.3681	1.3651	1.3631	1.3625
$MeNH_2$-MeOH	1.1387	1.1385	1.1382	1.1378	1.1374	1.1363	1.1350	1.1338	1.1333
$MeNH_2$-Peptide	1.1419	1.1415	1.1411	1.1407	1.1403	1.1390	1.1381	1.1376	1.1374
$MeNH_2$-Pyr	1.3394	1.3390	1.3386	1.3381	1.3377	1.3366	1.3356	1.3348	1.3344
MeOH-CH_3Cl	1.6106	1.6107	1.6106	1.6103	1.6098	1.6082	1.6066	1.6056	1.6051
MeOH-CH_3F	0.9805	0.9800	0.9794	0.9789	0.9784	0.9771	0.9758	0.9746	0.9740
MeOH-$MeNH_2$	1.3789	1.3795	1.3797	1.3795	1.3788	1.3757	1.3706	1.3657	1.3627
MeOH-MeOH	1.1446	1.1441	1.1435	1.1428	1.1420	1.1397	1.1371	1.1348	1.1333
MeOH-Peptide	1.1511	1.1501	1.1491	1.1481	1.1472	1.1446	1.1412	1.1382	1.1371
MeOH-Pyr	1.3493	1.3491	1.3487	1.3481	1.3472	1.3441	1.3401	1.3367	1.3344
MeOH-H_2O	1.1485	1.1482	1.1477	1.1470	1.1463	1.1444	1.1421	1.1401	1.1387
Peptide-$MeNH_2$	1.3768	1.3778	1.3781	1.3780	1.3775	1.3747	1.3701	1.3653	1.3622
Peptide-MeOH	1.1377	1.1374	1.1370	1.1367	1.1363	1.1354	1.1347	1.1342	1.1334
Peptide-H_2O	1.1439	1.1437	1.1433	1.1429	1.1425	1.1416	1.1407	1.1398	1.1388
H_2O-$MeNH_2$	1.3769	1.3777	1.3781	1.3779	1.3774	1.3745	1.3699	1.3655	1.3628
H_2O-MeOH	1.1443	1.1438	1.1433	1.1426	1.1418	1.1396	1.1370	1.1346	1.1332
H_2O-H_2O	1.1479	1.1477	1.1472	1.1466	1.1459	1.1440	1.1418	1.1397	1.1384

Table C.18: $\langle u \rangle$ for the n_Y LMO

System	$0.90d_0$	$0.95d_0$	$1.00d_0$	$1.05d_0$	$1.10d_0$	$1.25d_0$	$1.50d_0$	$2.00d_0$	∞
$AcNH_2$-$AcNH_2$	1.7297	1.7328	1.7349	1.7361	1.7367	1.7338	1.7411	1.7304	1.7015
AcOH-AcOH	1.7132	1.7202	1.7246	1.7274	1.7290	1.7297	1.7219	1.7286	1.6926
CCl_3OH-H_2O	1.7012	1.7059	1.7080	1.7085	1.7029	1.7004	1.6826	1.6549	1.6364
CF_3OH-H_2O	1.6893	1.6929	1.6946	1.6947	1.6939	1.6875	1.6728	1.6511	1.6365
HBr-MeOH	1.6912	1.6862	1.6822	1.6788	1.6759	1.6694	1.6623	1.6551	1.6519
HCl-$MeNH_2$	2.0303	2.0446	2.0533	2.0573	2.0573	2.0443	2.0181	1.9790	1.9486
HCl-MeOH	1.6844	1.6811	1.6783	1.6759	1.6738	1.6738	1.6618	1.6549	1.6517
HF-$MeNH_2$	1.9532	1.9696	1.9820	1.9912	1.9978	2.0064	1.9994	1.9697	1.9461
HF-MeOH	1.6677	1.6688	1.6694	1.6697	1.6695	1.6676	1.6625	1.6555	1.6517
$MeNH_2$-$MeNH_2$	1.9838	1.9835	1.9827	1.9814	1.9798	1.9740	1.9603	1.9521	1.9455
$MeNH_2$-MeOH	1.6842	1.6836	1.6827	1.6814	1.6799	1.6741	1.6646	1.6527	1.6503
$MeNH_2$-Peptide	1.7098	1.7094	1.7089	1.7083	1.7075	1.7022	1.6958	1.6934	1.6929
$MeNH_2$-Pyr	1.8931	1.8925	1.8918	1.8910	1.8900	1.8861	1.8803	1.8751	1.8737
MeOH-CH_3Cl	2.2416	2.2439	2.2449	2.2448	2.2439	2.2391	2.2336	2.2312	2.2292
MeOH-CH_3F	1.4479	1.4469	1.4456	1.4441	1.4423	1.4362	1.4290	1.4287	1.4178
MeOH-$MeNH_2$	1.9850	1.9912	1.9949	1.9968	1.9973	1.9939	1.9818	1.9580	1.9452
MeOH-MeOH	1.6869	1.6879	1.6879	1.6872	1.6861	1.6813	1.6710	1.6562	1.6498
MeOH-Peptide	1.7167	1.7196	1.7215	1.7224	1.7227	1.7207	1.7107	1.6949	1.6921
MeOH-Pyr	1.8922	1.8983	1.9026	1.9053	1.9070	1.9068	1.8992	1.8833	1.8737
MeOH-H_2O	1.6854	1.6852	1.6839	1.6820	1.6797	1.6724	1.6598	1.6429	1.6369
Peptide-$MeNH_2$	1.9927	1.9961	1.9978	1.9983	1.9979	1.9934	1.9806	1.9697	1.9445
Peptide-MeOH	1.6704	1.6693	1.6681	1.6668	1.6656	1.6620	1.6570	1.6518	1.6502
Peptide-H_2O	1.6670	1.6656	1.6640	1.6623	1.6606	1.6558	1.6496	1.6468	1.6369
H_2O-$MeNH_2$	1.9858	1.9911	1.9943	1.9959	1.9962	1.9928	1.9804	1.9563	1.9455
H_2O-MeOH	1.6861	1.6872	1.6872	1.6865	1.6854	1.6805	1.6699	1.6552	1.6496
H_2O-H_2O	1.6845	1.6844	1.6833	1.6815	1.6794	1.6721	1.6589	1.6414	1.6361

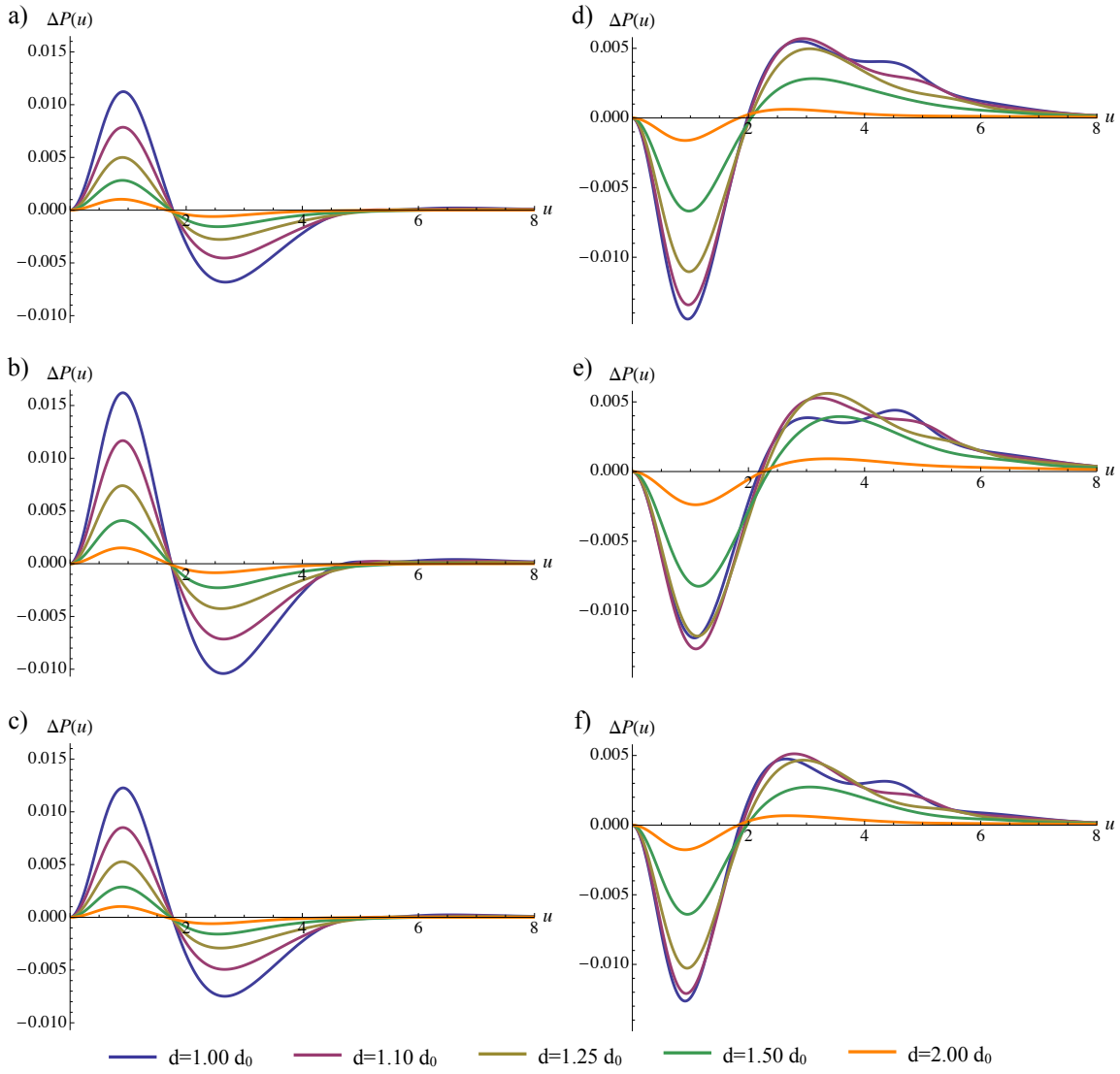


Figure C.1: $\Delta P(u)$ for the $\sigma_{\text{X-H}}$ bond LMOs of a) H_2O-H_2O, b) H_2O-$MeNH_2$, c) H_2O-MeOH, and the n_Y lone pairs of d) H_2O-H_2O, e) H_2O-$MeNH_2$, f) H_2O-MeOH.

Much like Figure 6.4, Figure C.1 demonstrates the effect of increasing the separation distance, d, between the hydrogen bond donor and acceptor species. The same decreasing effect is observed for these three species with H_2O as the donor species. Unsurprisingly, this trend is seen universally across the data set.

References

[1] Bernard, Y. A.; Loos, P. F.; Gill, P. M. W. *Mol. Phys.* **2013**, *111*, 2414–2426.

[2] *Bond Dissociation Energies in Simple Molecules*; U.S. National Bureau of Standards: Washington, 1970.

[3] Gordon, M. S.; Nguyen, K. A. *J. Phys. Chem.* **1989**, *93*, 7356–7358.

[4] Bento, A. P.; Sola, M.; Bickelhaupt, F. M. *J. Comput. Chem.* **2005**, *26*, 1497–1504.

[5] Knight, R. D. *Physics for Scientists and Engineers: A Strategic Approach*, 4th ed.; Pearson Education, Ltd.: New Jersey, 2016.

[6] Levine, I. N. *Quantum Chemistry*, 6th ed.; Pearson Education, Inc.: New Jersey, 2009.

[7] Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 1st ed.; John Wiley & Sons, Ltd.: West Sussex, 2004.

[8] Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education, Ltd.: Great Britain, 2001.

[9] Heisenberg, W. *Z. Phys.* **1927**, *43*, 172.

[10] Schrödinger, E. *Ann. Phys.* **1926**, *79*, 361.

[11] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: New Jersey, 1996.

[12] Born, M.; Oppenheimer, J. R. *Ann. Phys.* **1927**, *84*, 457.

[13] Hartree, D. R. *Proc. Camb. Phil. Soc.* **1927**, *24*, 89.

[14] Slater, J. C. *Phys. Rev.* **1930**, *35*, 210.

[15] Fock, V. *Z. Phys.* **1930**, *61*, 126.

[16] Slater, J. C.; Verma, H. C. *Phys. Rev.* **1929**, *34*, 1293.

[17] Slater, J. C. *Phys. Rev.* **1930**, *35*, 509.

[18] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

[19] Hall, G. G. *Proc. Royal Soc. (London)* **1951**, *A205*, 541.

[20] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbery, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[21] Shao, Y. et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

[22] Frisch, M. J. et al. GaussianâĹij09 Revision E.02. Gaussian Inc. Wallingford CT 2009.

[23] Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.

[24] Roothaan, C. C. J. *Rev. Mod. Phys.* **1960**, *32*, 179.

[25] Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.

[26] Boys, S. F. *Proc. Roy. Soc. (London)* **1950**, *A200*, 542–554.

[27] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

[28] Dunning, T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.

[29] Wilson, A. K.; Woon, D. E.; Peterson, K. E.; Dunning, T. H. *J. Chem. Phys.* **1999**, *110*, 7667–7676.

[30] Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 123.

[31] Weigend, F.; Alrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

[32] Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113.

[33] Jensen, F. *J. Chem. Phys.* **2002**, *116*, 3502.

[34] Jensen, F. *J. Chem. Phys.* **2004**, *121*, 3463.

[35] Jensen, F. *J. Phys. Chem. A* **2007**, *111*, 11198.

[36] Jensen, F. *J. Chem. Phys.* **2012**, *136*, 114107.

[37] Chandrasekhar, J.; Andrade, J. G.; Schleyer, P. V. *J. Am. Chem. Soc.* **1981**, *103*, 5609.

[38] Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1509.

[39] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

[40] Sherill, C. D.; Schaefer, H. F. In *Advances in Quantum Chemistry*; Löwdin, P. O., Ed.; Academic Press: San Diego, 1999; pp 143–269.

[41] Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

[42] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

[43] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

[44] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[45] Schultz, N.; *et al.*, *J. Phys. Chem. A* **2005**, *109*, 4388.

[46] Schultz, N.; *et al.*, *J. Phys. Chem. A* **2005**, *109*, 11127.

[47] Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415.

[48] Vosko, S. H.; WIlk, L.; Nusair, M. *Can. J., Phys.* **1980**, *58*, 1200.

[49] Becke, A. D. *Phys. Rev.* **1988**, *A38*, 3098.

[50] Becke, A. D. *J. Chem. Phys.* **1992**, *96*, 2155.

[51] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1986**, *33*, 8800.

[52] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

[53] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[54] Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

[55] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.

[56] Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.

[57] Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

[58] Coulson, C. A.; Neilson, A. H. *Proc. Phys. Soc. London* **1961**, *78*, 831.

[59] Gill, P. M. W.; Crittenden, D. L.; O'Neill, D. P.; Besley, N. A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 15–25.

[60] Dumont, E. E.; Crittenden, D. L.; Gill, P. M. W. *Phys. Chem. Chem. Phy.* **2007**, *9*, 5340.

[61] Crittenden, D. L.; Dumont, E. E.; Gill, P. M. W. *J. Chem. Phys.* **2007**, *127*, 141103.

[62] Bernard, Y. A.; Crittenden, D. L.; Gill, P. M. W. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3447.

[63] Pearson, J. K.; Crittenden, D. L. *J. Chem. Phys.* **2009**, *130*, 164110.

[64] Crittenden, D. L.; Gill, P. M. W. In *Solving the Schrödinger Equation: Has Everything Been Tried*; Popelier, P. L. A., Ed.; IC Press: London, 2011; pp 1–23.

[65] Proud, A. J.; Pearson, J. K. *J. Chem. Phys.* **2010**, *133*, 134113.

[66] Shepard, B. J.; Pearson, J. K. *Theor. Chem. Acc.* **2017**, *136*.

[67] Banyard, K. E.; Reed, C. E. *J. Phys. B* **1978**, *11*, 2957.

[68] Rassolov, V. A. *J. Chem. Phys.* **1999**, *110*, 3672.

[69] Bernard, Y. A.; Gill, P. M. W. *New J. Phys.* **2009**, *11*, 083015.

[70] Bernard, Y. A.; Gill, P. M. W. *J. Phys. Chem. Lett.* **2010**, *1*, 1254–1258.

[71] Bernard, Y. A.; Crittenden, D. L.; Gill, P. M. W. *J. Phys. Chem. A* **2010**, *114*, 11984–11991.

[72] Wigner, E. *Phys. Rev.* **1932**, *40*, 749–659.

[73] Hollett, J. M.; Gill, P. M. W. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2972–2978.

[74] Coleman, A. J. *Int. J. Quantum Chem.* **1967**, *1*, 457.

[75] Thakkar, A. J.; Moore, N. J. *Int. J. Quantum Chem.* **1981**, *20*, 393–400.

[76] Galvez, F. J.; Buendia, E.; Sarsa, A. *Journal of Chemical Physics* **1999**, *111*, 3319.

[77] Lennard-Jones, J. E.; Pople, J. A. *Proc. Roy. Soc. (London)* **1950**, *A202*, 166.

[78] Lennard-Jones, J. E.; Pople, J. A. *Proc. Roy. Soc. (London)* **1950**, *A210*, 190.

[79] Edmiston, C.; Ruedenberg, K. *Rev. Mod. Phys.* **1963**, *35*, 457–464.

[80] Foster, J.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[81] Pipek, J.; Mezey, P. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[82] Zielinski, Z. A. M.; Pearson, J. K. *Comp. Theor. Chem.* **2013**, *1003*, 79–90.

[83] Hennessey, D. C.; Sheppard, B. J. H.; Mackenzie, D. E. C. K.; Pearson, J. K. *Phys. Chem. Chem. Phys.* **2014**, *16*, 25548.

[84] Lewis, G. N. *J. Am. Chem. Soc.* **1916**, *38*, 762–785.

[85] *Valence and the Structure of Atoms and Molecules*, 1st ed.; The Chemical Catalog Company, Inc.: New York, 1923.

[86] Gillespie, R. J.; Nyholm, R. S. *Quart. Rev. Chem. Soc.* **1957**, *11*, 339–380.

[87] Gillespie, R. J. *J. Chem. Educ.* **1970**, *47*, 18–23.

[88] Weinhold, F.; Klein, R. A. *Mol. Phys.* **2012**, *110*, 565.

[89] Weinhold, F.; Klein, R. A. *Chem. Educ. Res. Pract.* **2014**, *15*, 276.

[90] Bruice, P. Y. *Organic Chemistry*, 7th ed.; Pearson Education, Ltd.: New Jersey, 2012.

[91] Proud, A. J.; Mackenzie, D. E. C. K.; Pearson, J. K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 20194.

[92] Proud, A. J.; Pearson, J. K. *Can. J. Chem.* **2016**, *94*, 1077.

[93] Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.

[94] Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.

[95] Weinhold, F.; Carpenter, J. E. In *The Structure of Small Molecules and Ions*; Naaman, R., Vager, Z., Eds.; Plenum: New York, 1988; pp 227–236.

[96] Glendening, E. D.; Landis, C. R.; Weinhold, F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 1–42.

[97] Weinhold, F. *J. Comp. Chem.* **2012**, *33*, 2363–2379.

[98] Becke, A. D.; E., E. K. *J. Chem. Phys.* **1990**, *92*, 5397–5403.

[99] Bader, R. W. F. *Atoms in Molecules: A Quantum Theory*; Oxford University Press, 1994.

[100] Sarasola, C.; Ugalde, J. M.; Boyd, R. J. *J. Phys. B: At. Mol. Opt. Phys.* **1990**, *23*, 1095.

[101] Benesch, R.; Smith, V. H. *Acta. Crystallogr. Sect. A* **1970**, *26*, 586.

[102] Boyd, R. J.; Coulson, C. A. *J. Phys. B* **1973**, *6*, 782.

[103] Thakkar, A. J.; Smith, V. H. *Chem. Phys. Lett.* **1976**, *42*, 782.

[104] Boyd, R. J.; Yee, M. C. *J. chem. Phys.* **1982**, *77*, 3578.

[105] Thakkar, A. J.; Tripathi, A. N.; Smith, V. H. *Int. J. Quantum Chem.* **1984**, *26*, 157.

[106] Thakkar, A. J. In *Density Matrices and Density Functionals*; Erdahl, R. M., Smith, V. H., Jr., Eds.; Reidel: Dordrecht, Holland, 1987; pp 553–581.

[107] Boyd, R. J.; Sarasola, C.; Ugalde, J. M. *J. Phys. B* **1988**, *21*, 2555.

[108] Ugalde, J. M.; Sarasola, C.; Dominguez, L.; Boyd, R. J. *J. Math. Chem.* **1991**, *6*, 51.

[109] Wang, J.; Smith, V. H. *Int. J. Quantum Chem.* **1994**, *49*, 147.

[110] Cioslowski, J.; Stefanov, B.; Tang, A.; Umrigar, C. J. *J. Chem. Phys.* **1995**, *103*, 6093.

[111] Cioslowski, J.; Liu, G. *J. Chem. Phys.* **1996**, *105*, 4151.

[112] Cioslowski, J.; Liu, G. *J. Chem. Phys.* **1993**, *109*, 8225.

[113] Fradera, X.; Sarasola, C.; Ugalde, J. M.; Boyd, R. J. *Chem. Phys. Lett.* **1999**, *304*, 393.

[114] Lee, A. M.; Gill, P. M. W. *Chem. Phys. Lett.* **1999**, *313*, 271.

[115] Valderrama, E.; Fradera, X.; Ugalde, J. M. *J. Chem. Phys.* **2001**, *115*, 1987.

[116] Valderrama, E.; Ugalde, J. M. *Int. J. Quantum Chem.* **2002**, *86*, 40.

[117] Koga, T.; Matsuyama, H.; Dehesa, J. S.; Thakkar, A. J. *J. Chem. Phys.* **1999**, *110*, 5763.

[118] Matsuyama, H.; Koga, T. *Chem. Phys. Lett.* **1999**, *300*, 515.

[119] Koga, T. *Chem. Phys. Lett.* **2001**, *350*, 135.

[120] Gill, P. M.; O'Neill, D. P.; Besley, N. A. *Theoretical Chemistry Accounts* **2003**, *109*, 241–250.

[121] Reed, C. E.; Banyard, K. E. *J. Phys. B* **1980**, *13*, 1519.

[122] Mobbs, R. J.; Banyard, K. E. *J. Chem. Phys.* **1983**, *78*, 6106.

[123] Levin, V. G.; Neudatchin, V. G.; Pavlitchenkov, A. V.; Smirnov, Y. F. *J. Phys. B* **1984**, *17*, 1525.

[124] Ugalde, J. M. *J. Phys. B* **1987**, *20*, 2153.

[125] Youngman, P. K.; Banyard, K. E. *J. Phys. B* **1987**, *20*, 3313.

[126] Banyard, K. E.; Youngman, P. K. *J. Phys. B* **1987**, *20*, 5585.

[127] Banyard, K. E.; Al-Bayati, K. H.; Youngman, P. K. *J. Phys. B* **1988**, *21*, 3177.

[128] Banyard, K. E.; Mobbs, R. J. *J. Chem. Phys.* **1988**, *88*, 3788.

[129] Banyard, K. E.; Sanders, J. *J. Chem. Phys.* **1993**, *99*, 5281.

[130] Thakkar, A. J. *Chem. Phys. Lett.* **2003**, *381*, 80–85.

[131] Wang, J.; Smith, V. H. *J. Chem. Phys.* **1993**, *99*, 9745.

[132] Keeble, D. R. T.; Banyard, K. E. *J. Phys. B* **1997**, *30*, 13.

[133] Koga, T.; Matsuyama, H. *J. Chem. Phys.* **1997**, *107*, 8510.

[134] Matsuyama, H.; Koga, T.; Romera, E.; Dehesa, J. S. *Phys. Rev. A* **1998**, *57*, 1759.

[135] Koga, T.; Matsuyama, H. *J. Chem. Phys.* **1998**, *111*, 643.

[136] Matsuyama, H.; Koga, T.; Kato, Y. *J. Phys. B* **1999**, *32*, 3371.

[137] Koga, T.; Matsuyama, H. *J. Chem. Phys.* **2000**, *113*, 10114.

[138] Koga, T.; Kato, Y.; Matsuyama, H. *Theor. Chem. Acc.* **2001**, *106*, 237.

[139] Koga, T. *J. Chem. Phys.* **2002**, *116*, 6614.

[140] Koga, T. *Theor. Chem. Acc.* **2002**, *107*, 246.

[141] Besley, N. A.; Lee, A. M.; Gill, P. M. *Molecular Physics* **2002**, *100*, 1763–1770.

[142] Dominguez, L.; Aguado, M.; Sarasola, C.; Ugalde, J. M. *J. Phys. B: At. Mol. Opt. Phys.* **1992**, *25*, 1137.

[143] Wang, J.; Tripathi, A. N.; Jr, V. H. S. *J. Phys. B: At. Mol. Opt. Phys.* **1993**, *26*, 205.

[144] Koga, T.; Matsuyama, H.; Romera, E.; Dehesa, J. S. *Phys. Rev. A* **1998**, *57*, 4212.

[145] Koga, T.; Matsuyama, H. *J. Chem. Phys.* **1998**, *108*, 3424.

[146] Galvez, F. J.; Buendia, E.; Sarsa, A. *J. Chem. Phys.* **1999**, *111*, 3319.

[147] Koga, T.; Matsuyama, H.; Molina, J. M.; Dehesa, J. S. *Eur. Phys. J. D* **1999**, *7*, 17.

[148] McCarthy, S. P.; Thakkar, A. J. *Int. J. Quantum Chem.* **2011**, *111*, 753.

[149] Stowasser, R.; Hoffmann, R. *J. Am. Chem. Soc.* **1999**, *121*, 3414–3420.

[150] Mura, M.; Knowles, P. *J. Chem. Phys.* **1996**, *104*, 9848–9858.

[151] Řezáč, J.; Riley, K. E.; Hobza, P. *Journal of Chemical Theory and Computation* **2011**, *7*, 2427–2438.

[152] Grabowski, S. J. *J. Phys. Chem. A* **2001**, *105*, 10739.

[153] Parthasarathi, R.; Subramanian, V.; Sathyamurthy, N. *J. Phys. Chem. A Lett.* **2006**, *110*, 3349.

[154] Clark, T.; Hennemann, M.; Murray, J. S.; Politzer, P. *J. Mol. Model.* **2007**, *13*, 291.

[155] Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y.; Murray, J. S. *J. Mol. Model.* **2007**, *13*, 305.

[156] Mathematica 8. Wolfram Research, Inc., 2010.

[157] Mahaffy, P. G.; Bucat, B.; Tasker, R.; Kotz, J. C.; Treichel, P. M.; Weaver, G. C.; McMurry, J. *Chemistry: Human Activity, Chemical Ractivity*, 1st ed.; Nelson Education, Ltd.: Toronto, 2011.

[158] Koga, T.; Matsuyama, H. *J. Phys. B: At. Mol. Opt. Phys.* **1997**, *30*, 5631.

[159] Koga, T. *J. Chem. Phys.* **1998**, *108*, 2515.

[160] Mercero, J. M.; Fowler, J.; Sarasola, C.; Ugalde, J. M. *Phys. Rev. A* **1999**, *59*, 4255.

[161] Valderrama, E.; Fradera, X.; Ugalde, J. M. *J. Chem. Phys.* **2001**, *115*, 1987.

[162] Valderrama, E.; Mercero, J. M.; Ugalde, J. M. *J. Phys. B: At. Mol. Opt. Phys.* **2001**, *34*, 275.

[163] de Meijere, A. *Angew. Chem. Int. Ed.* **1979**, *18*, 809–826.

[164] Coulson, C. A.; Moffitt, W. E. *Phil. Mag.* **1949**, *40*, 1.

[165] Bernett, W. A. *J. Chem. Educ.* **1967**, *44*, 17.

[166] Ayers, P. L. et al. *Comp. Theor. Chem.* **2015**, *1053*, 2–16.

[167] Jensen, W. B. *J. Chem. Educ.* **1996**, *73*, 11–20.

[168] Jensen, W. B. *J. Chem. Educ.* **2003**, *80*, 279–287.

[169] Pauling, L. *J. Am. Chem. Soc.* **1932**, *54*, 3570–3582.

[170] Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782–784.

[171] Gordy, W. *Phys. Rev.* **1946**, *69*, 604–607.

[172] Walsh, A. D. *Proc. R. Soc., London* **1951**, *A207*, 13–22.

[173] Huggins, M. L. *J. Am. Chem. Soc.* **1953**, *75*, 4123–4126.

[174] Sanderson, R. T. *J. Chem. Phys.* **1955**, *23*, 2467.

[175] Allen, A. L. *J. Inorg. Nuc. Chem.* **1958**, *5*, 264.

[176] Hinze, J.; Jaffe, H. H. *J. Am. Chem. Soc.* **1962**, *84*, 540–546.

[177] Hinze, J.; Whitehead, M. A.; Jaffe, H. H. *J. Am. Chem. Soc.* **1963**, *85*, 148–154.

[178] Phillips, J. C. *Phys. Rev. Lett.* **1968**, *20*, 550–553.

[179] Michaelson, H. B. *IBM J. Res. Dev.* **1978**, *22*, 72–80.

[180] Allen, L. C. *J. Am. Chem. Soc.* **1992**, *144*, 1510–1511.

[181] Simons, G.; Zandler, M. E.; Talaty, E. R. *J. Am. Chem. Soc.* **1976**, *98*, 7869–7870.

[182] Boyd, R. J.; Edgecombe, K. E. *J. Am. Chem. Soc.* **1988**, *110*, 4182–4186.

[183] Boyd, R. J.; Boyd, S. L. *J. Am. Chem. Soc.* **1992**, *114*, 1652–1655.

[184] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 17b, September 2015, Editor: Russell D. Johnson III http://cccbdb.nist.gov/.

[185] Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: Ithaca, 1939.

[186] Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theor. Comput.* **2012**, *8*, 4285–4292.

[187] Popelier, P. L. A. *J. Phys. Chem. A* **1998**, *102*, 1873.

[188] Wojtulewski, S.; Grabowski, S. J. *J. Chem. Phys.* **2005**, *309*, 183.

[189] Zierkiewicz, W.; Jurecka, P.; Hobza, P. *Chem. Phys. Chem.* **2005**, *6*, 609.

[190] Gora, R. W.; Grabowski, S. J.; Leszczynski, J. *J. Phys. Chem. A* **2005**, *109*, 6397.