A Density Functional Odyssey Beyond Ground State Energies

by

Diptarka Hait

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Martin Head-Gordon, Chair
Professor Eric Neuscamman
Professor Jeffrey Neaton

Spring 2022

A Density Functional Odyssey Beyond Ground State Energies

Copyright 2022
by
Diptarka Hait

Abstract

A Density Functional Odyssey Beyond Ground State Energies

by

Diptarka Hait

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Martin Head-Gordon, Chair

Density functional theory or DFT is presently the most popular route for computing the electronic structure of chemical systems. Although DFT is formally exact, the exact functional that maps the electron density to the energy remains unknown to date. A large number of density functional approximations (DFAs) have consequently been developed to compute the energies of molecules and extended materials. Use of exact constraints, large amounts of highly accurate benchmark data, and intelligent data-driven design schemes have resulted in modern functionals that provide an excellent balance between computational cost and predictive accuracy. However, nearly all the DFA development efforts in recent years had focused on improving chemically relevant energy differences in the ground state. Even the electron density, which is the central quantity of the theory, has been mostly neglected. This dissertation tries to explore usage of DFT beyond ground state energies through the investigation of electrical response properties and electronic excited states. Information from these regimes should prove helpful extending the applicability of DFT beyond computation of ground state energy differences, and also assist in designing more transferable DFAs that better approximate the exact functional.

The first half of the dissertation assesses the accuracy of modern DFAs in predicting molecular properties associated with the response of the energy to electric fields. The exact functional is formally capable of predicting exact energies even when the system is subjected to arbitrary electric fields. However, approximate functionals only model the electrical response well if the density is accurate. The ability of DFAs to compute electrical response properties thus indicate their accuracy in modeling densities. The dissertation therefore studies dipole moments (Chapter 4), second cumulants of the density (Chapter 5) and static dipole polarizabilities (Chapter 6). High level coupled cluster benchmarks at the complete basis set limit for ≥ 100 chemical species has been generated for all the three properties. These benchmark datasets are used to evaluate the performance of several popular and recent functionals, in order to gauge performance. This analysis reveals that some of the most accurate modern

DFAs for ground state energies yield suboptimal predictions for electrical response properties. Future DFA development therefore should utilize these benchmark datasets for training and assessment purposes, in order to obtain functionals that simultaneously yield accurate energies and densities. In addition, we use the static dipole polarizability as a sensitive probe for electronic structure in Chapter 7, to identify qualitative problems in DFAs. This demonstrates that several modern DFAs are challenged by homolytic single bond dissociation, as they fail to completely unpair electrons over the right distance scales. The material in this half of the dissertation therefore provides information about how existing functionals struggle to model density, and should be helpful for the design of more accurate DFAs.

The second half of the dissertation examines behavior for electronic excited states, focusing on the popular linear-response time-dependent DFT (TDDFT) and the less well known orbital optimized DFT (OO-DFT) approaches. Chapter 10 shows that TDDFT methods cannot describe bond dissociations in the excited state, developing unphysical derivative discontinuities at the onset of spin unpairing in the ground state. The other chapters focus on OO-DFT, and applications to core spectroscopy. Chapter 11 presents a robust new algorithm for excited state OO, that ensures the optimization process remains on the chosen state and does not undergo ‘variational collapse’, to a lower energy state. This Square Gradient Minimization or SGM algorithm is used to model core-level excitations for closed-shell systems (Chapter 12) and radicals (Chapter 13), using OO-DFT. Chapter 13 also presents a scheme to recouple three unpaired electrons to obtain spin-pure doublets, which are relevant for core to unoccupied orbital transitions in radicals. The results of Chapter 12 and 13 demonstrate the OO-DFT with the SCAN DFA can model core-level spectra of second period elements to semiquantitative accuracy of ~ 0.3 eV, against experimental values with ~ 0.1 eV uncertainty. This is a dramatic improvement over the ~ 15 eV errors observed from TDDFT, indicating that OO-DFT/SCAN is a cheap and reliable way to model core-level spectra. Indeed, OO-DFT/SCAN can be directly used to simulate experimental spectra, such as time-resolved X-ray transient absorption studies of chemical dynamics. The energies and densities of these core-excited states also provide new information for functional training beyond ground state energies. Incorporation of this very distinct form of data in the DFA development process thus can help better approximate the exact functional.

Ὁ βίος βραχύς,	Life is short,
ἡ δὲ τέχνη μακρή,	and art long,
ὁ δὲ καιρὸς ὀξύς,	the crisis fleeting,
ἡ δὲ πεῖρα σφαλερή,	experience perilous,
ἡ δὲ κρίσις χαλεπή.	and decision difficult.

Hipprocrates. “Αφορισμοί”. Translated by Francis Adams (1849).

Humankind cannot gain anything without first giving something in return. To obtain, something of equal value must be lost.
Hiromu Arakawa. “Fullmetal Alchemist”.

Contents

Contents **ii**

1 Overview of Quantum Chemistry and Density Functional Theory **1**
- 1.1 Basics of Quantum Chemistry . . . 1
  - 1.1.1 The Hamiltonian . . . 1
  - 1.1.2 Units . . . 3
  - 1.1.3 The Hydrogen Atom . . . 3
  - 1.1.4 General Chemical Systems . . . 3
  - 1.1.5 Slater Determinant . . . 5
  - 1.1.6 Hartree-Fock Theory . . . 6
  - 1.1.7 Wavefunction Theories . . . 9
  - 1.1.8 Density Matrices . . . 10
- 1.2 Density Functional Theory . . . 11
  - 1.2.1 Background . . . 11
  - 1.2.2 The Hohenberg-Kohn Theorems . . . 12
  - 1.2.3 Kohn-Sham Theory . . . 14
  - 1.2.4 Local Functionals . . . 16
    - 1.2.4.1 Local Density Approximation . . . 16
    - 1.2.4.2 Generalized Gradient Approximation . . . 18
    - 1.2.4.3 Meta-GGAs . . . 18
  - 1.2.5 Delocalization Error . . . 19
  - 1.2.6 Hybrid Functionals . . . 21
  - 1.2.7 Double Hybrids . . . 23
  - 1.2.8 Jacob's Ladder . . . 23
  - 1.2.9 Dispersion Interactions . . . 25
  - 1.2.10 Spin Symmetry Breaking . . . 26
  - 1.2.11 Beyond Ground State Energies . . . 27

2 Outline of the Dissertation **29**
- 2.1 Works Not Discussed in This Dissertation . . . 35

I Electrical Response Properties 37

3 Foreword to Electrical Response Properties 38
- 3.1 The Density Problem 38
- 3.2 Connecting Densities to Energy 39
- 3.3 Multipole Moments 40
- 3.4 Multipole Polarizabilities 42
- 3.5 Objectives 43

4 Dipole Moments 44
- 4.1 Introduction 44
- 4.2 Computational Methods 45
- 4.3 Results for Equilibrium Geometries 47
  - 4.3.1 Performance of Wavefunction Theory 47
  - 4.3.2 Performance of DFT 50
  - 4.3.3 Basis Set Convergence 52
- 4.4 Dipole Moments at Non-equilibrium Configurations 54
- 4.5 Conclusion 57
- 4.6 Supporting Information 58

5 Second Cumulants of the Electron Density 59
- 5.1 Introduction 59
- 5.2 Second Spatial Cumulant of Electron Density 60
  - 5.2.1 Definition 60
  - 5.2.2 Connection to Multipole Moments 61
  - 5.2.3 Physical Interpretation 62
  - 5.2.4 Behavior vs System Size 62
- 5.3 Dataset 65
  - 5.3.1 Computational Details 66
- 5.4 Results and Discussions 67
  - 5.4.1 Full Dataset 67
  - 5.4.2 The Case of Challenging Atoms 70
- 5.5 Stretched H_2 72
- 5.6 Discussion and Conclusion 73
- 5.7 Supporting Information 74

6 Static Polarizabilities 75
- 6.1 Introduction 75
- 6.2 Computational Methods 77
- 6.3 Results 79
  - 6.3.1 Performance of Wavefunction Theory 79
  - 6.3.2 Performance of DFT 82

6.3.3 Challenging Species . . . 84
6.3.4 Basis Set Convergence . . . 86
6.4 Conclusion . . . 87
6.5 Supporting Information . . . 88

7 Polarizabilities beyond the Coulson-Fischer Point 89
7.1 Introduction . . . 89
7.2 Computational Methods . . . 92
7.3 Results and Discussions . . . 92
7.3.1 Qualitatively Correct Functionals . . . 92
7.3.2 Static Polarizabilities . . . 94
7.3.3 Force Constants . . . 97
7.3.4 Relation to Lowest Eigenvalue of Orbital Rotation Hessian . . . 98
7.3.5 Spin Localization . . . 100
7.3.6 Density Corrected Polarizabilities . . . 104
7.3.7 Hybrid Functionals . . . 105
7.3.8 Barriers to H Atom Association . . . 106
7.4 Discussion and Conclusions . . . 107
7.5 Appendix: Orbital Rotation Hessian . . . 108
7.6 Supporting Information . . . 109

8 Afterword to Electrical Response Properties 110

II Electronic Excited States 112

9 Foreword to Electronic Excited States 113
9.1 Linear Response Time-Dependent DFT . . . 114
9.2 Orbital Optimized DFT . . . 116
9.2.1 Variational Collapse . . . 119
9.2.2 Multiconfigurational States . . . 120
9.2.3 Formal Justification . . . 121
9.3 Core-level Spectroscopy . . . 122
9.4 Objectives . . . 122

10 TDDFT beyond the Coulson-Fischer Point 124
10.1 Introduction . . . 124
10.2 Spin-flipping Excitations within TDDFT/TDHF . . . 126
10.3 CIS and TDHF for Stretched H_2 in a Minimal Basis. . . . 128
10.4 Stretched H_2 in a Larger Basis. . . . 133
10.5 Examples for Other Systems . . . 138
10.5.1 $M_S = 0$ subspace . . . 138

10.5.2 $M_S = \pm 1$ subspaces 138
10.5.3 RCIS 141
10.6 Conclusion 143
10.7 Computational Details 144
10.8 Appendix: Spin-flip Block TDHF Solutions for Minimal Basis H_2 144
10.9 Supporting Information 145

11 Square Gradient Minimization 146
11.1 Introduction 146
11.2 Theory 147
11.2.1 Objective Function 147
11.2.2 Gradient 148
11.2.3 Preconditioner 149
11.2.4 The SGM Method for Orbital Optimization 149
11.2.5 Relationship with Other Methods 150
11.2.6 Local Extrema in Δ 151
11.3 Excited State Orbital Optimized DFT 151
11.3.1 ΔSCF 151
11.3.2 ROKS 152
11.3.3 Constrained Methods 153
11.4 Applications 153
11.4.1 Comparison of SGM to MOM and IMOM 153
11.4.2 Application to Doubly Excited States 154
11.4.3 Singly Excited States of Formaldehyde 157
11.5 Low Energy Excited States of Zinc Phthalocyanine 159
11.6 Summary and Discussion 161
11.7 Computational Details 162
11.8 Supporting Information 162

12 Restricted Open-Shell Kohn-Sham for Core-Level Excitations 163
12.1 Main Text 163
12.2 Appendix: Spin-orbit effects in L-edge spectra 176
12.3 Computational Details 177
12.4 Supporting Information 177

13 Core-Level Spectra of Open-Shell Systems 178
13.1 Introduction 178
13.2 Theory 180
13.2.1 Single Determinant States 180
13.2.2 Multiconfigurational States 181
13.2.3 Transition Dipole Moments 183
13.3 Results and Discussion 184

    - 13.3.1 Excitations to the SOMO . . . 184
        - 13.3.1.1 The Case of NH_2^+ . . . 188
    - 13.3.2 Spectrum of the Allyl Radical . . . 189
    - 13.3.3 O K-edge Spectrum of CO^+ . . . 192
    - 13.3.4 N K-edge Spectrum of NO_2 . . . 194
- 13.4 Recommendations for Successful Calculations . . . 195
- 13.5 Conclusions . . . 197
- 13.6 Computational Methods . . . 198
- 13.7 Appendix: Phase Estimation for Mixed Configurations . . . 198
- 13.8 Supporting Information . . . 199

14 Afterword to Electronic Excited States **200**

Bibliography **204**

Acknowledgments

This dissertation has only one listed author, but the truth is that it only became possible through the efforts of many. The first and foremost of those is my advisor, Prof. Martin Head-Gordon. There are legions of academically better qualified people who can wax eloquent about how brilliant Martin is as a scientist, how good his knowledge/insight is, etc. Personally though, what stands out the most is his fundamental decency as both an advisor and a person. I first met him as a college junior, and was immediately struck by how nice and friendly this respected senior academic was. Seven years have passed since then (six spent in his group), and my first impression remains unaltered. Thanks a lot for all the advice over the years, for giving me the freedom to pursue my interests, for always helping, and for your willingness to change your mind when presented with evidence. Also, sorry for all the typos you had to fix, and all the letters I made you write. I hope that at least some of your positive way of thinking about science and life, have percolated into my being.

I also thank my other dissertation committee members, Profs. Eric Neuscamman and Jeff Neaton for taking time out of their busy schedule go over this leviathan of a document, and for the constructive suggestions. In addition, I really appreciate all the conversations I had with Eric about science and life in academia over the years, not to mention the letters.

A large part of my core-level spectroscopy work was done in close contact with Prof. Steve Leone's group. They taught me what questions spectroscopists were interested in and where theory could really help; greatly expanding my initial naive perspective. Thanks for accepting me into your community and for your patience with my slowness in both comprehension and computation. In particular, I'd like to thank Steve for all the advice, support, and letters I got from him. They were invaluable.

My journey into physical chemistry began with a quantum mechanics class taught by Bob Field and Troy Van Voorhis back in freshman year. I then had Troy as my research advisor and Bob as my academic advisor for the next three years. I owe a great deal of my graduate school successes to their mentorship in my formative years as a scientist. Thank you for always having your doors open for a chat (even for an ignorant freshman).

My coworkers in the MHG group have played a key role in shaping me as a scientist. In particular, I'd like to thank Dan for being a fantastic summer mentor, who made sure I didn't sink in my first month and has always been willing to lend a ear in the years since. Chatting random stuff with you is always a pleasure, and I am really grateful for your patience when I am being about slow about science. I'd also like to thank Norm for the infectious enthusiasm for all things electronic structure in general and for getting me into selected CI in particular. I think we have had a quite productive partnership that taught me a bunch of new things. Yuezhi, thanks for being such a supportive senior group member and for all the Q-Chem support. Zhou, thanks for being a fantastic colleague over two groups, and all the advice over the years. Adam, thanks for being my first (and most prolific!) mentee and for making sure I go to the gym regularly. Two chapters in this dissertation arose from our partnership, and I learned a lot about how to be a better mentor from working with you. Jason (Liang), thanks for working on the cumulants with me, and for the boundless enthusiasm about science. Good luck at Columbia, you'll do fantastic and I'll be paying attention from the other shore! Leo, thanks for messing with relativity together, I do hope you find the happiness you are seeking with the bears. Richard (Kang), thanks for being a fantastic coworker and you'll do great in grad school. I'm only sorry that I will no longer have a front row ticket to witness it. Jiashu, Josie, Katherine, James, Matthias, Luke, Abdul, Brad, Srimukh, Juan and Alex: thanks for the camaraderie and for all the cool science we have discussed together. Last but certainly not the least, thanks Jessie for handling all the administrative things and for making sure I was paid!

My social circle in Berkeley consisted of many of the aforementioned, and their presence kept me afloat through graduate school. Beyond the MHG group, I'd like to thank Andrew, Song, Carlos, Chin, Jason (Fang), and Nathan for the company over the years. I've done many fun activities with overlapping subsets of you: hikes, food, anime, stealing rides, and just plain hanging out. My life has been enriched by your presence-thank you for being a part of my odyssey. And I certainly won't complain about the research partnerships that arose in the process.

A special shoutout to my fellow beavers from college, whose friendship continues to remain a major part of my life. Lily, I don't know what I'd have done without all our chats over the years, to remind me there are things outside the bubble of theoretical chemistry that matter much more. Or without your proofreading skills, for that matter. If I have any regrets about this moment, it is that we didn't get to graduate together. Angus, thanks for being a great friend even as I ghosted you to write this dissertation. I loved our adventures, and hope we can hit up some more national parks! Tomo, James, and Aofei, thanks for being such great friends who add an element of levity to life. Charles, thanks for making sure I had a roof above my head when I first came to Berkeley, and for always including me in activities, whether boardgames or hikes. Hongzhou, thanks for being a great conversation partner over the years. Not to mention how much your research inspired parts of this document.

My research was financially supported by the Department of Energy, through the Office of Science, Office of Basic Energy Sciences. I was personally supported early in graduate school by a Berkeley Fellowship, and at the very end by a Pimentel Research Award (endowed by Prof. Lester Andrews). I am very grateful to all organizations and people that supported my research interests, and hope that I have been able to make good use of it.

I'd also like to thank my family for their continuous, and unconditional support. Perhaps it would have been easier if you folks weren't looking up my papers on arXiv all the time to pester me about what journal they have been submitted to, but I suppose I'm not used to having it any other way. A special thanks to the first (and only other) immediate family PhD, my uncle Dr. Arghya Hait. Thanks for getting me into science when I was a kid, for sustaining that interest over the years, for making me aware what a PhD was, and for being my first (and ongoing) role model. Finally, thanks to my parents for making sure I got to this moment. Neither of you got a chance to reach this far, and I'm sorry that the best I can do is just let you experience it vicariously through me.

Chapter 1

Overview of Quantum Chemistry and Density Functional Theory

We do not expect an accurate description of chemical binding.
W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects." *Phys. Rev.,* **140**, A 1138, 1965.

Quantum mechanics (QM) is used to describe the behavior of matter and energy on an atomic scale [1]. QM is therefore essential for understanding chemistry at the molecular level. This is however not a trivial task, as the interactions between nuclei and electrons in most chemical species lead to a great deal of complexity that is extremely difficult to model [2]. Closed-form solutions have only been found for a few model systems, and numerical solutions are computationally unfeasible for all but the smallest chemical species. The discipline of quantum chemistry therefore seeks to develop computationally affordable approximations to exact QM, that have reasonable predictive utility for chemical systems.

1.1 Basics of Quantum Chemistry

1.1.1 The Hamiltonian

QM modeling usually begins with solving the time-independent Schrödinger equation. Each system has an associated Hamiltonian operator $\mathbf{H}$, leading to an eigenproblem of the form:

$$\mathbf{H}\left|\psi_k\right\rangle = E_k\left|\psi_k\right\rangle \tag{1.1}$$

The "stationary" states of the system are described by $\{\left|\psi_k\right\rangle\}$, with energies $\{E_k\}$ (k being the state label). Any arbitrary state of the system can be expressed as a linear combination of the eigenstates $\{\left|\psi_k\right\rangle\}$, indicating that solving Eq 1.1 is critical.

Within nonrelativistic QM and in the absence of external fields, $\mathbf{H}$ for chemical systems can generally be expressed as follows:

$$\mathbf{H} = \mathbf{T}_e + \mathbf{T}_n + \mathbf{V}_{ne} + \mathbf{V}_{nn} + \mathbf{V}_{ee} \tag{1.2}$$

$\mathbf{T}_e$ and $\mathbf{T}_n$ are the kinetic energies of the electrons and nuclei respectively. $\mathbf{V}_{ne}$,$\mathbf{V}_{nn}$ and $\mathbf{V}_{ee}$ are purely electrostatic (Coulombic) terms that describe electron-nuclear attraction, nuclear-nuclear repulsion and electron-electron repulsion respectively.

It is worth recalling that the nuclei are (at least) several thousand times heavier than electrons, and therefore are relatively more "classical" in behavior. The quantum chemistry problem can therefore be further simplified by decoupling the electronic and nuclear degrees of freedom from each other. This is achieved by the Born-Oppenheimer approximation [3, 4], which treats the nuclei as classical point charges. Only the electronic degrees of freedom are treated quantum mechanically, leading to a new Hamiltonian:

$$\mathbf{H}_e(\{\vec{R}\}) = \mathbf{T}_e + \mathbf{V}_{ne}(\{\vec{R}\}) + \mathbf{V}_{ee} + V_{nn}(\{\vec{R}\}) \tag{1.3}$$

where $\{\vec{R}\}$ are the (fixed) positions of the nuclei. This approximation is therefore equivalent to setting the kinetic energy of the nuclei $\mathbf{T}_n$ in Eq 1.2 to zero, and the nuclear-nuclear repulsion operator $\mathbf{V}_{nn}$ to the scalar valued function $V_{nn}(\{\vec{R}\})$ that describes classical electrostatic repulsion between point charge nuclei at positions $\{\vec{R}\}$.

Electronic structure theory aims to solve this electronic Hamiltonian $\mathbf{H}_e(\{\vec{R}\})$ via obtaining energies $E_{e,k}(\{\vec{R}\})$ and states $\left|\psi_k(\{\vec{R}\})\right\rangle$ that satisfy:

$$\mathbf{H}_e(\{\vec{R}\})\left|\psi_k(\{\vec{R}\})\right\rangle = E_{e,k}(\{\vec{R}\})\left|\psi_k(\{\vec{R}\})\right\rangle \tag{1.4}$$

This protocol is generally adequate for describing most chemical problems. If desired, nuclear QM effects can subsequently be well approximated by solving the nuclear Schrödinger equation described by the Hamiltonian:

$$\mathbf{H}_n = \mathbf{T}_n + E_e(\{\vec{\mathbf{R}}\}) \tag{1.5}$$

where the effective potential energy operator $E_{e,k}(\{\vec{\mathbf{R}}\})$ is the operator form of the electronic energy function $E_{e,k}(\{\vec{R}\})$.

The Born-Oppenheimer approximation therefore leads to a picture of molecules as nuclei confined within potential energy surfaces $E_{e,k}(\{\vec{R}\})$ produced by the electrons, which in popular parlance is often simplified to "balls (nuclei) connected by springs (that represent the energy surface)". This decoupling of nuclear and electronic degrees of freedom via the Born-Oppenheimer approximation is applicable to a broad range of chemistry, but breaks down when the energy surfaces $E_{e,k}(\{\vec{R}\})$ corresponding to different electronic states are close in energy (i.e. states with different k are near-degenerate for some $\{\vec{R}\}$) due to the failure of the adiabatic approximation [5]. We will however not consider this regime further, but instead direct interested readers to Ref 6 for further details. In subsequent sections therefore, we will assume static nuclei, drop explicit references to $\{\vec{R}\}$ for states, and only refer to the electronic Hamiltonian as $\mathbf{H}_e$.

1.1.2 Units

Throughout this work, we will use atomic units unless specified otherwise. Specifically, we will operate in a system of units where $\hbar$, $4\pi\epsilon_0$, the mass of an electron m_e, and the elementary charge e are all defined to be 1.

1.1.3 The Hydrogen Atom

The hydrogen atom is the simplest possible quantum chemical system, and is one of the few with an analytic solution. $\mathbf{H}_e$ for this problem in the spatial representation is:

$$\mathbf{T}_e = -\frac{1}{2}\nabla^2; \mathbf{V}_{ne} = -\frac{1}{r} \tag{1.6}$$

$$\mathbf{H}_e = \mathbf{T}_e + \mathbf{V}_{ne} \tag{1.7}$$

where the nucleus is placed at the origin, and $\vec{r}$ is the position of the electron.

The spatial representation ('wavefunction') of the eigenstates $|nlm\rangle$ are [7]:

$$\langle\vec{r}|nlm\rangle = \phi_{nlm}(\vec{r}) = \sqrt{\left(\frac{2}{n}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} e^{-\frac{r}{n}} r^l L_{n-l-1}^{2l+1}\left(\frac{r}{n}\right) Y_l^m(\theta,\phi) \tag{1.8}$$

where r, θ, ϕ are the polar coordinates corresponding to the vector $\vec{r}$. L is a Laguerre polynomial, while Y is a spherical harmonic function. The quantum numbers $n, l, m \in \mathbb{Z}$ describe the structure of the solutions, with $n > l \geq |m| \geq 0$. It is worth noting that the wavefunction very far from the nucleus ($r \to \infty$) is exponentially decaying in r. In general, the radial distribution of the electron in space is provided via $e^{-\frac{r}{n}} r^l L_{n-l-1}^{2l+1}\left(\frac{r}{n}\right)$ while $Y_l^m(\theta,\phi)$ gives the angular distribution. $\phi_{nlm}(\vec{r})$ are often described as 'orbitals', originating from the incorrect classical view about the electron 'orbiting' the nucleus like a miniature solar system.

While $\phi_{nlm}(\vec{r})$ is fairly complicated, the energy eigenvalues are simply given by:

$$E_{nlm} = -\frac{1}{2n^2} \tag{1.9}$$

Interestingly, this n^{-2} form was empirically found by Rydberg in the late nineteenth century [8], utilizing experimental data about spectral lines of atomic hydrogen. Eqn 1.9 is therefore a beautiful example of the synergy between experiment and theory that continues to drive modern chemical physics. In particular, it highlights the very strong connection between atomic/molecular spectroscopy and QM theory that continues to date.

1.1.4 General Chemical Systems

For a general molecule composed of m electrons and M nuclei (with indices $\{i, j \ldots\}$ and $\{A, B \ldots\}$ respectively), the molecular Hamiltonian in the spatial representation is typically given by:

$$\mathbf{T}_e = -\frac{1}{2}\sum_i \nabla_i^2 \tag{1.10}$$

$$\mathbf{V}_{ne} = -\sum_{i,A} \frac{Z_A}{\left|\vec{R}_A - \vec{r}_i\right|} \tag{1.11}$$

$$\mathbf{V}_{ee} = \frac{1}{2}\sum_{i,j;i\neq j} \frac{1}{\left|\vec{r}_i - \vec{r}_j\right|} \tag{1.12}$$

$$V_{nn} = \frac{1}{2}\sum_{A,B;A\neq B} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|} \tag{1.13}$$

$$\mathbf{H}_e = \mathbf{T}_e + \mathbf{V}_{ne} + \mathbf{V}_{ee} + V_{nn} \tag{1.14}$$

where $\{\vec{r}_i\}$ and $\{\vec{R}_A\}$ are the positions of the electrons and nuclei respectively, while $\{Z_A\}$ are the nuclear charges. Aside from the presence of the interelectron repulsion term $\mathbf{V}_{ee}$, the overall form looks quite similar to that of the hydrogen atom (as V_{nn} is a scalar). It therefore seems possible to utilize products of single particle orbitals $\phi_i(\vec{r}_i)$ to describe the full many-electron wavefunction $\Psi(\{\vec{r}_i\})$.

However, this form of $\mathbf{H}_e$ obscures two important aspects. Eqn 1.14 gives the impression that all the electrons can be distinguished from each other. On the other hand, we know that electrons are indistinguishable quantum particles, and therefore exchanging any two electrons should not change any physical observables [5]. The Stern-Gerlach experiment [9] also reveals that aside from spatial position $\vec{r}_i$, there is an additional degree of freedom per electron called spin s_i, which is an intrinsic angular momentum for quantum particles and is not at all referenced in Eqn 1.14.

These additional factors are accounted by Fermi-Dirac (FD) statistics [10, 11], which requires that the exchange of any two electrons should only result in changing the sign of the wavefunction. In other words, for a general wavefunction $\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)$ FD statistics states that swapping electrons 1 and 2 should lead to:

$$\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots) = -\Psi(\vec{r}_2, s_2; \vec{r}_1, s_1; \vec{r}_3, s_3 \ldots) \tag{1.15}$$

The many-electron wavefunction is therefore antisymmetric to the pairwise exchange of electrons. This constraint needs to be directly enforced on any attempt to solve Eqn 1.14.

One interesting consequence is that no two electrons can have the same position and spin coordinates. This is easily shown by noting that if electrons 1 and 2 have the same position $\vec{r}$ and spin s, we'd have from Eqn 1.15:

$$\Psi(\vec{r}, s; \vec{r}, s; \vec{r}_3, s_3 \ldots) = -\Psi(\vec{r}, s; \vec{r}, s; \vec{r}_3, s_3 \ldots) = 0 \tag{1.16}$$

This is colloquially referred to as the Pauli exclusion principle [12]: "Two electrons can't simultaneously occupy the same quantum state".

1.1.5 Slater Determinant

The simplest way to ensure FD statistics is by solving Eqn 1.14 only within the subspace spanned by antisymmetric states (i.e. states that change signs upon exchange of any two electrons). Given a set of electrons $\{1, 2 \ldots m\}$ that occupy spin-orbitals $\{\phi_i, \phi_j \ldots\}$, a simple and computationally efficient antisymmetric wavefunction is the Slater determinant [13]:

$$\Phi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots) = \frac{1}{\sqrt{m!}} \begin{vmatrix} \phi_i(\vec{r}_1, s_1) & \phi_i(\vec{r}_2, s_2) & \phi_i(\vec{r}_3, s_3) \ldots \phi_i(\vec{r}_m, s_m) & \\ \phi_j(\vec{r}_1, s_1) & \phi_j(\vec{r}_2, s_2) & \phi_j(\vec{r}_3, s_3) \ldots \phi_j(\vec{r}_m, s_m) & \\ \vdots & \vdots & \ddots & \vdots \end{vmatrix} \tag{1.17}$$

in the position-spin representation. The orbitals $\{\phi_i, \phi_j \ldots\}$ are single-particle wavefunctions that can be expressed as:

$$\phi_i(\vec{r}, s) = c_\alpha \chi_i^\alpha(\vec{r})\alpha(s) + c_\beta \chi_i^\beta(\vec{r})\beta(s); c_\alpha, c_\beta \in \mathbb{C} \tag{1.18}$$

where α and β represent eigenfunctions for the spin up and down states respectively. More precisely:

$$\mathbf{S}_z \left|\alpha\right\rangle = \frac{1}{2} \left|\alpha\right\rangle \quad \text{and} \quad \mathbf{S}_z \left|\beta\right\rangle = -\frac{1}{2} \left|\beta\right\rangle \tag{1.19}$$

where $\mathbf{S}_z$ is the z component of the spin operator $\vec{\mathbf{S}}$.

$\chi_i^\alpha(\vec{r})$ and $\chi_i^\beta(\vec{r})$ are referred to as spatial orbitals as they only depend on $\vec{r}$. This '(spin)-general' form is however not very widely used because the resulting expressions are rather complicated. A more common representation is:

$$\phi_i(\vec{r}, s) = \chi_i^\alpha(\vec{r})\alpha(s) \quad \text{or} \quad \chi_i^\beta(\vec{r})\beta(s) \tag{1.20}$$

which assigns a single spin to an orbital, making a great deal of the mathematics easier (typically with no degradation in results). This form is referred to as '(spin)-unrestricted'. A more constrained form is the '(spin)-restricted' case where $\chi_i^\beta(\vec{r}) = \chi_i^\alpha(\vec{r}) \,\forall i$.

As an example of a restricted Slater determinant, let us consider the lowest energy state (ground state) of the helium atom. The lowest energy orbital of the hydrogen atom is $\psi_{100}(\vec{r})$, which is referred to as the 1s level. If we assume that the helium atom has a similar 1s spatial orbital $\chi_{1s}(\vec{r})$, containing one up and one down spin, we have:

$$\phi_i(\vec{r}, s) = \chi_{1s}(\vec{r})\alpha(s) \quad \text{and} \quad \phi_j(\vec{r}, s) = \chi_{1s}(\vec{r})\beta(s) \tag{1.21}$$

$$\Phi(\vec{r}_1, s_1; \vec{r}_2, s_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_i(\vec{r}_1, s_1) & \phi_i(\vec{r}_2, s_2) \\ \phi_j(\vec{r}_1, s_1) & \phi_j(\vec{r}_2, s_2) \end{vmatrix} \tag{1.22}$$

$$= \frac{1}{\sqrt{2}} \chi_{1s}(\vec{r}_1)\chi_{1s}(\vec{r}_2) \left(\alpha(s_1)\beta(s_2) - \beta(s_1)\alpha(s_2)\right) \tag{1.23}$$

We note that this factorized form was only possible because both orbitals had the same spatial component $\chi_{1s}(\vec{r})$. In general, spin-restricted Slater determinants are factorizable into spatial and spin components. More general spin-unrestricted and spin-general determinants need not have this property. As an example, we can consider the spin-unrestricted case of two infinitely separated hydrogen atoms A and B, with one up and one down spin. Both of these spins will occupy 1s orbitals χ_{1sA} and χ_{1sB} centered on different atoms A and B. Therefore:

$$\phi_i(\vec{r}, s) = \chi_{1sA}(\vec{r})\alpha(s) \tag{1.24}$$

$$\phi_j(\vec{r}, s) = \chi_{1sB}(\vec{r})\beta(s) \tag{1.25}$$

$$\Phi(\vec{r}_1, s_1; \vec{r}_2, s_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_i(\vec{r}_1, s_1) & \phi_i(\vec{r}_2, s_2) \\ \phi_j(\vec{r}_1, s_1) & \phi_j(\vec{r}_2, s_2) \end{vmatrix} \tag{1.26}$$

$$= \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_{1sA}(\vec{r}_1)\alpha(s_1) & \chi_{1sA}(\vec{r}_2)\alpha(s_2) \\ \chi_{1sB}(\vec{r}_1)\beta(s_1) & \chi_{1sB}(\vec{r}_2)\beta(s_2) \end{vmatrix} \tag{1.27}$$

$$= \frac{1}{\sqrt{2}} \left(\chi_{1sA}(\vec{r}_1)\alpha(s_1)\chi_{1sB}(\vec{r}_2)\beta(s_2) - \chi_{1sB}(\vec{r}_1)\beta(s_1)\chi_{1sA}(\vec{r}_2)\alpha(s_2)\right) \tag{1.28}$$

which cannot be further simplified or factorized.

This has some interesting consequences, as $\mathbf{H}_e$ has no explicit dependence on $\vec{\mathbf{S}}$. In particular, the squared spin magnitude operator $\mathbf{S}^2$ and $\mathbf{H}_e$ commute (i.e. $\mathbf{H}_e\mathbf{S}^2 - \mathbf{S}^2\mathbf{H}_e = 0$), indicating that an eigenstate of $\mathbf{H}_e$ must be an eigenstate of $\mathbf{S}^2$. Exact wavefunctions therefore are factorizable into separate spatial and spin components. Spin-unrestricted and spin-general Slater determinants need not have this property and thus are not guaranteed to be eigenstates of $\mathbf{S}^2$ (unlike spin-restricted determinants). Nonetheless, they can be very useful approximations to the true wavefunction despite this lack of spin symmetry, as will be shown later. In particular, a system with equal number of up and down spins can always be represented by a restricted determinant. However, there may exist spin symmetry broken spin-unrestricted or spin-general determinants of lower energy. The energies of these spin symmetry broken determinants are better approximations to the exact ground state energy (through the variational theorem [14]). Ref 15 provides a detailed discussion on the stability of Slater determinants.

1.1.6 Hartree-Fock Theory

Let us consider a Slater determinant $|\Phi\rangle$ of m electrons occupying orthonormal spin-orbitals $\{\phi_i, \phi_j \ldots\}$. The expectation value of the energy (E_0) is given by:

$$E_0 = \langle\Phi| \mathbf{H}_e |\Phi\rangle = \langle\Phi| \mathbf{T}_e |\Phi\rangle + \langle\Phi| \mathbf{V}_{ne} |\Phi\rangle + \langle\Phi| \mathbf{V}_{ee} |\Phi\rangle + V_{nn} \tag{1.29}$$

$$V_{nn} = \frac{1}{2} \sum_{A,B;A\neq B} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|} \tag{1.30}$$

$$\langle\Phi| \mathbf{T}_e |\Phi\rangle = -\frac{1}{2} \sum_i \int \phi_i^*(\vec{r}, s)\nabla^2\phi_i(\vec{r}, s)d\vec{r}ds \tag{1.31}$$

$$\langle\Phi|\,\mathbf{V}_{ne}\,|\Phi\rangle = -\sum_i \int \phi_i^*(\vec{r},s)\sum_A \frac{Z_A}{\left|\vec{R}_A-\vec{r}\right|}\phi_i(\vec{r},s)d\vec{r}ds \tag{1.32}$$

$$\langle\Phi|\,\mathbf{V}_{ee}\,|\Phi\rangle = \frac{1}{2}\sum_{i,j;i\neq j}(J_{ij}-K_{ij}) \tag{1.33}$$

$$J_{ij} = \int \phi_i^*(\vec{r}_1,s_1)\phi_j^*(\vec{r}_2,s_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\phi_i(\vec{r}_1,s_1)\phi_j(\vec{r}_2,s_2)d\vec{r}_1ds_1d\vec{r}_2ds_2 \tag{1.34}$$

$$K_{ij} = \int \phi_i^*(\vec{r}_1,s_1)\phi_j^*(\vec{r}_2,s_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\phi_j(\vec{r}_1,s_1)\phi_i(\vec{r}_2,s_2)d\vec{r}_1ds_1d\vec{r}_2ds_2 \tag{1.35}$$

V_{nn} is just the scalar corresponding to the nuclear-nuclear electrostatic repulsion energy. $\langle\Phi|\,\mathbf{T}_e\,|\Phi\rangle$ and $\langle\Phi|\,\mathbf{V}_{ne}\,|\Phi\rangle$ are sums over the expectation value of these operators for each orbital. J_{ij} represents the average electrostatic repulsion between an electron in orbital i with another in orbital j. K_{ij} however eludes a simple explanation, and is a purely quantum mechanical term arising from FD statistics reducing electron-electron repulsion. More precisely, Pauli exclusion prevents two electrons of the same spin from closely approaching each other in space and therefore the mean-field repulsion J_{ij} has to be reduced by K_{ij}. The spin connection can be seen very clearly for the case of spin-unrestricted orbitals, where $K_{ij}=0$ unless ϕ_i and ϕ_j have the same spin, indicating that only repulsion between electrons of the same spin is reduced by this term. K is therefore referred to as the 'exchange' contribution, since it arises from FD statistics regarding exchange of electrons.

The energy E_0 can be variationally minimized through optimizing the orbitals $\{\phi_i\}$. This protocol yields the single best Slater determinant approximation to the true ground state wavefunction and is called the Hartree-Fock (HF) method [14]. HF is often referred to as a mean-field theory, because the electrons experience the average field of all other electrons (with the sum $\frac{1}{2}\sum\limits_{ij,i\neq j} J_{ij}$ being the classical electrostatic self-energy of the electrons). The exchange terms make this interpretation slightly less straightforward, but it remains quite reasonable to view HF as a theory that considers interelectron interactions in a mean-field manner, to the extent that it is possible within FD statistics. We note that Hartree's original model [16] lacked exchange (as it preceded knowledge about FD statistics) and was thus quite inaccurate. It was only corrected later by Fock [17] and Slater [18] to the modern HF model.

One interesting thing to note is that $J_{ii}=K_{ii}$ and so we can state:

$$\langle\Phi|\,\mathbf{V}_{ee}\,|\Phi\rangle = \frac{1}{2}\sum_{i,j;i\neq j}(J_{ij}-K_{ij}) = \frac{1}{2}\sum_{i,j}(J_{ij}-K_{ij}) \tag{1.36}$$

On the surface, inclusion of J_{ii} does not make any sense as an electron in orbital ϕ_i should not repel itself. However, this exact cancellation between 'self-repulsion' and 'self-exchange' in HF theory is a powerful feature, that will subsequently become relevant as many other QM approximations lack this property.

HF is quite computationally efficient, scaling as the fourth power as the molecular size N (i.e. an asymptotic computational scaling of $O(N^4)$) for most implementations targeting molecular systems. However, it is also quite incomplete due to the mean-field treatment of interelectron repulsion (especially for opposite spins). For example, Eqn 1.23 indicates that the joint probability density for the two electrons in the helium atom is:

$$p(\vec{r}_1, \vec{r}_2) = \int |\Phi(\vec{r}_1, s_1; \vec{r}_2, s_2)|^2 \, ds_1 ds_2 = |\chi_{1s}(\vec{r}_1)\chi_{1s}(\vec{r}_2)|^2 = |\chi_{1s}(\vec{r}_1)|^2 \, |\chi_{1s}(\vec{r}_2)|^2 \tag{1.37}$$

which indicates that the two electrons are spatially independent of each other and thus uncorrelated. This appears to be somewhat absurd, since electrons should try to avoid each other. Indeed, the spherical symmetry of χ_{1s} ensures that $p(\vec{r}, \vec{r}) = p(\vec{r}, -\vec{r})$, which suggests that the probability of both electrons of being at the same spot is identical to them being diametrically opposite each other. This unphysical uncorrelated behavior of electrons leads to an overestimation of the electron-electron repulsion energy of the helium atom with HF (as compared to exact quantum mechanics). Consequently, the HF estimate for first ionization energy of the helium atom is 1 eV lower than experiment [7]. A similar behavior is seen for dihydrogen, where the bond dissociation energy evaluated by HF is 3 eV vs 4 eV from exact QM or experiment [7]. This is problematic, as a chemical bond is central to chemistry and 1 eV$\sim$ 23 kcal/mol $\sim$ 12000 K, making it an unacceptable level of error. HF theory is therefore incapable of yielding quantitative accuracy for most chemical problems, despite its conceptual simplicity and computational efficiency. The gap between HF and exact QM energies is termed as 'correlation energy' since it represents the contribution that is missed on account of restricting the wavefunction to a single Slater determinant (which amounts to an effective mean-field treatment of interelectron interactions).

Before proceeding further, it seems reasonable to estimate the size of all of these individual energy subcomponents for a representative chemical system, in order to gauge their relative importance. For a water molecule in the electronic ground state at equilibrium (minimum energy) geometry, we have at the HF level:

1. Kinetic energy: $\langle\Phi|\,\mathbf{T}_e\,|\Phi\rangle = 2068.0$ eV
2. Electron-nucleus attraction:$\langle\Phi|\,\mathbf{V}_{ne}\,|\Phi\rangle =$ -5416.5 eV
3. Classical interelectron repulsion: $\frac{1}{2}\sum\limits_{i,j} J_{ij}$ =1272.0 eV (includes self-repulsion J_{ii})
4. Exchange: $-\frac{1}{2}\sum\limits_{i,j} K_{ij}$ =-243.4 eV (includes self-exchange terms K_{ii})
5. Internuclear repulsion: $V_{nn} = 250.1$ eV

In addition, the correlation energy is $\sim$ -10.1 eV.

V_{nn} is a scalar, and thus of little consequence. Considering the other terms, we see that electron-nucleus attraction is the largest in magnitude, followed by kinetic energy and classical repulsion. The exchange term is much smaller than interelectron repulsion, while correlation is the smallest by magnitude. This small value however belies its critical importance for chemical modeling, which relies on *relative energies* (i.e. energy differences) and not absolute energies. Chemical processes generally entail formation or cleavage of chemical bonds, which in turn entails pairing or unpairing of electrons. Each valence electron pair roughly has $\sim$ 1 eV associated correlation energy, and a perceptible change of correlation energy is generally associated with chemical processes. In addition, dispersion interactions require correlation between induced dipoles for energy lowering, indicating that correlation energy plays a central role in describing noncovalent interactions.

1.1.7 Wavefunction Theories

Exact quantitative modeling of electron correlation can be achieved by variational optimization of the wavefunction within the Hilbert space spanned by all possible Slater determinants. This protocol is described as full configuration interaction (FCI) in the literature, as each Slater determinant corresponds to a different electronic configuration. FCI however has a computational cost that scales exponentially with system size, and cannot be applied to systems bigger than ethane routinely even with modern computers.

The task of modeling correlation becomes much easier if the exact FCI wavefunction is dominated by a single Slater determinant. HF is qualitatively acceptable in this 'single reference' limit, and the remaining 'weak'/'dynamic' electron correlation can be well approximated by perturbative or coupled cluster approaches as corrections to the mean-field HF reference. We will not discuss these approaches in detail here, but instead direct the interested reader to Ref 19 for the former and Ref 20 for the latter (as well as the references therein). We only note that the simplest perturbative model is second-order Møller-Plesset theory [21] (MP2), which scales as the fifth power of the molecular size N (i.e. $O(N^5)$ asymptotic computational scaling). Similarly, the most popular coupled cluster method is coupled cluster singles and doubles (CCSD), which scales as $O(N^6)$. Both MP2 and CCSD represent a significant improvement over HF, but do not generally attain the widely accepted 'chemical accuracy' limit of 1 kcal/mol [22]. The latter is only reliably achieved by inclusion of perturbative triples into CCSD, with the resulting CCSD(T) method [23] popularly being labeled as the 'gold standard' of quantum chemistry.

There are also systems where certain electrons experience considerable interelectron repulsion, as compared to the sum of their kinetic energy and nuclear attraction. The relatively high magnitude of repulsion compared to attractive energy terms cause such electrons to generally be amongst the highest in energy. This significant level of electron-electron repulsion leads to unpairing of electron pairs, and thereby results in multiple Slater determinants having significant weight in the FCI wavefunction. These systems are therefore called 'multireference'/'strongly correlated' and are difficult to model with the single reference methods discussed previously. HF tends to undergo spin symmetry breaking for such systems in an attempt to unpair electrons, but the resulting spin-polarized Slater determinants are not eigenstates of $\mathbf{S}^2$.

The most reliable approach to model such species is to perform FCI within a subspace (‘active space’) of orbitals consisting of electrons that are expected to be strongly interacting/correlated. The limit for doing this exactly is $\sim$ 24 electrons in $\sim$ 24 orbitals with existing modern supercomputers [24]. Ref 25 and references therein describe approximate FCI solvers that can handle larger active spaces going up to 30 electrons in 108 orbitals to approximately acheive chemical accuracy. Such methods however remain exponentially scaling, and are impractical for any system larger than a benzene molecule.

1.1.8 Density Matrices

Most of the computational complexity in the preceding section arises from the sheer size of the Hilbert space available to the electrons. Each electron has 4 degrees of freedom (3 spatial, 1 spin), leading to the wavefunction existing in a $4m$ dimensional space for m electrons. This appears to be somewhat unnecessary, as electrons interact at most pairwise. Indeed, given a wavefunction $\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)$ we know that the interelectron repulsion is:

$$\frac{1}{2}\sum_{i,j;i\neq j}\int \Psi^*(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)\frac{1}{|\vec{r}_i-\vec{r}_j|}\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)d\vec{r}_1 ds_1 d\vec{r}_2 ds_2 d\vec{r}_3 ds_3 \ldots \tag{1.38}$$

$$= \frac{m(m-1)}{2}\int \Psi^*(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)\frac{1}{|\vec{r}_1-\vec{r}_2|}\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)d\vec{r}_1 ds_1 d\vec{r}_2 ds_2 d\vec{r}_3 ds_3 \ldots \tag{1.39}$$

$$= \int \frac{1}{|\vec{r}_1-\vec{r}_2|}G(\vec{r}_1, s_1; \vec{r}_2, s_2)d\vec{r}_1 ds_1 d\vec{r}_2 ds_2 \tag{1.40}$$

where

$$G(\vec{r}_1, s_1; \vec{r}_2, s_2) = \frac{m(m-1)}{2}\int \Psi^*(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)\Psi(\vec{r}_1, s_1; \vec{r}_2, s_2; \vec{r}_3, s_3 \ldots)d\vec{r}_3 ds_3 \ldots \tag{1.41}$$

The simplification arises from the indistinguishibility of the electrons (which makes each term of the sum in Eqn 1.38 identical). $G(\vec{r}_1, s_1; \vec{r}_2, s_2)$ is related to the joint probability density of finding two electrons: one at $\vec{r}_1$ with spin s_1 and the other at $\vec{r}_2$ with spin s_2. $G(\vec{r}_1, s_1; \vec{r}_2, s_2)$ is in fact the representation of a density operator called the two-particle reduced density matrix (2RDM) or Γ, which is defined as follows for a given electronic state $|\Psi\rangle$:

$$\Gamma = \frac{m(m-1)}{2}\underset{3,4,\ldots n}{\mathbf{Tr}}\left[|\Psi\rangle\langle\Psi|\right] \tag{1.42}$$

wherein the full density operator $|\Psi\rangle\langle\Psi|$ has its trace taken over all but two electrons (i.e. electrons $3, 4 \ldots m$) and only has two electronic degrees of freedom left. It can be shown that the total electronic energy can obtained as a simple closed-form expression in terms of Γ without requiring any further information about the state $|\Psi\rangle$ [26]. It is therefore only necessary to know the total pairwise distribution of electrons instead of keeping track of all of them [27]. Of course, finding Γ without use of $|\Psi\rangle$ is nontrivial, as any arbitrary matrix would not correspond to a m electron antisymmetric wavefunction [28]. This so called 'N-representability' problem has defied accurate and efficient solutions. In particular, it is necessary to constrain higher order reduced density matrices to be positive semidefinite [29] to ensure an N-representable Γ. This is quite computationally demanding [30] and only a subset of such constraints [27, 31, 32] are usually employed by approximate 2RDM theories in practice [33, 34].

The 2RDM approach also raises the question as to whether further simplification is possible. Specifically, is it possible that the electronic energy can be found from nothing more than the electron number density alone? The answer is yes, with the resulting family of methods being referred to as density functional theory (DFT).

1.2 Density Functional Theory

DFT aims to map the electron number density function $\rho(\vec{r})$ (henceforth referred to as density) to the ground state energy E (with the word 'functional' describing a map that converts a function to a scalar). This idea is computationally appealing, as it indicates that the total energy of the system can be found from a function of 3 spatial variables ($\rho(\vec{r})$), instead of a wavefunction spanning the entirety of an exponential scaling Hilbert space.

1.2.1 Background

The idea behind DFT dates back to Thomas [35], Fermi [36] and Dirac [37] (TFD) in the 1920s. The TFD model decomposes E into the kinetic energy T, electron-nuclear attraction E_{ne}, electron-electron repulsion E_J, exchange E_x, correlation E_c and V_{nn}. In other words:

$$E = T + E_{ne} + E_J + E_x + E_c + V_{nn} \tag{1.43}$$

From classical electrostatics, E_{ne} and E_J are exactly:

$$E_{ne} = -\int \rho(\vec{r}) \left(\sum_A \frac{Z_A}{\left|\vec{R}_A - \vec{r}\right|} \right) d\vec{r} \tag{1.44}$$

$$E_J = \frac{1}{2} \int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2 \tag{1.45}$$

Note that E_J contains self-repulsion, as it is merely the classical self-energy of a continuous charge distribution (here $\rho(\vec{r})$ from the electron cloud). From HF results on the uniform electron gas (UEG), we can estimate T and E_x to be:

$$T = \frac{3(3\pi^2)^{\frac{2}{3}}}{10} \int (\rho(\vec{r}))^{5/3}\, d\vec{r} \tag{1.46}$$

$$E_x = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int (\rho(\vec{r}))^{4/3}\, d\vec{r} \tag{1.47}$$

No estimates for E_c were known at the time and it was thus neglected. As we shall see, the TFD model faced far greater problems than this.

The principal challenge with the TFD approach was the T functional, for which Eqn 1.46 was exact only for the UEG. Most molecular systems are quite removed from the uniform density limit, with rapidly varying density in and around the nucleus. The magnitude of T typically far exceeds E_x or E_c. Therefore, inaccuracies in the T functional in practice turned out to be a bigger issue than very precisely modeling E_x or E_c. TFD kinetic energies deviate from HF by 5-10% for atoms [38], which is unacceptably large. From Sec 1.1.6, we see that 5% error in T for H_2O amounts to $\sim$ 100 eV, while correlation is $\sim$ 10 eV. More general models for T (such as the von Weizsäcker correction [39]) did not adequately alleviate the problem. This doomed the practical utility of the TFD model for molecular systems. Indeed, it was shown that TFD failed to predict the very existence of a chemical bond [40]! In contrast, HF does not have such issues and is able to yield a qualitatively accurate description of electronic structure in almost all cases, including a reasonable description of bonding.

There are two more subtle issues with the TFD approach as stated above. The first is that it was not known at the time whether two different ground state wavefunctions (for two different Hamiltonians) could lead to the same $\rho(\vec{r})$. The functional can only map the eigenenergy to one scalar, and thus the presence of two distinct eigenstates with the same density would drastically limit the applicability of such an approach.

The second problem is that the exact ground state density is not known in advance. Any protocol would therefore have to find $\rho(\vec{r})$ as well. It is tempting to suggest that the optimal $\rho(\vec{r})$ should minimize the energy functional. However, it remained unproven at the time whether the ground state density $\rho(\vec{r})$ truly minimized the energy functional or not. Furthermore, it is essential that $\rho(\vec{r})$ arises from a proper, antisymmetric wavefunction of m electrons (i.e. be N-representable). The challenge of DFT is thus to know whether there exists a functional that can map $\rho \to E$, and how to find the exact ground state ρ.

1.2.2 The Hohenberg-Kohn Theorems

In 1964, Hohenberg and Kohn (HK) presented a formally exact variational principle [41] for E in terms of ρ. Specifically, they proved two theorems for an arbitrary system of m interacting electrons subjected to some external local potential $v_{ext}(\vec{r})$:

1. $v_{ext}(\vec{r})$ is uniquely specified (to an additive constant) by $\rho(\vec{r})$. This means the system Hamiltonian can be uniquely determined from $\rho(\vec{r})$, and so can E (via solving the Schrödinger equation). Therefore, let the unique functional that maps $\rho \to E$ be $E[\rho]$.

2. $E[\rho]$ has its minimum value for the true ground state density ρ. In other words, the density that minimizes the functional $E[\rho]$ is the true ground state density.

Since $v_{ext}(\vec{r}) = -\sum_A \frac{Z_A}{\left|\vec{R}_A - \vec{r}\right|}$ represents the electrostatic potential generated by the nuclei within the Born-Oppenheimer approximation, the HK theorems apply to the electronic Hamiltonian $\mathbf{H_e}$. They are however more general and can be used with other well-behaved local potentials $v_{ext}(\vec{r})$, including chemical systems subjected to arbitrary electric fields. Notably, the HK theorems do not directly apply to cases involving nonlocal [42] potentials, but we do not consider such systems in this work.

We will not present the full proof of the HK theorems, but instead direct interested readers to the original work (Ref 41). We will however present an intuitive 'proof' offered by E. Bright Wilson [43, 44] for the first HK theorem as applied to $\mathbf{H_e}$. Given some ground state $\rho(\vec{r})$ we can find the total number of electrons $m = \int \rho(\vec{r})d\vec{r}$ quite easily. The nuclear positions $\{\vec{R}_A\}$ are also specified by 'cusps' in the total density in space (i.e. points where the density has a first derivative discontinuity due to a change of sign). Finally, the nuclear charges $\{\vec{Z}_A\}$ can be found via:

$$\left(\frac{1}{2\rho(\vec{r})}\frac{d\rho(\vec{r})}{d\vec{r}}\right)_{\vec{r}\to\vec{R}_A} = -Z_A \tag{1.48}$$

This is known as Kato's cusp condition [45], and can be intuitively understood by noting that the Coulomb potential from the nuclei is singular at position of any given nucleus, and so the local behavior there can be estimated by ground state hydrogen atom like expressions (via Eqn 1.9) as all other potential energy terms are relatively insignificant in magnitude. $\mathbf{H_e}$ is thus completely specified by $\rho(\vec{r})$ as the total electron number m, nuclear positions $\{\vec{R}_A\}$ and charges $\{\vec{Z}_A\}$ can be directly determined from it. However, this argument cannot accommodate the presence of non-singular external potentials, such as a constant electric field. Nonetheless, it offers an intuitive 'feel' of how the HK theorems might work.

We note that the original HK proof was restricted to nondegenerate ground states and implicitly assumed that $\rho(\vec{r})$ arises from the ground state of some $v_{ext}(\vec{r})$ (i.e. is 'v-representable'). A constrained search formalism defined by Levy [46] removes both of these limitations, by defining a universal functional $\mathcal{Q}[\rho]$ given by:

$$\mathcal{Q}[\rho] = \min_{|\Psi_\rho\rangle\to\rho} \langle\Psi_\rho|\, \mathbf{T}_e + \mathbf{V}_{ee}\, |\Psi_\rho\rangle \tag{1.49}$$

In other words, $\mathcal{Q}[\rho]$ for some $\rho(\vec{r})$ is the minimum possible $\langle\Psi_\rho|\, \mathbf{T}_e + \mathbf{V}_{ee}\, |\Psi_\rho\rangle$ over the set of all possible many-electron wavefunctions $|\Psi_\rho\rangle$, that are constrained to have $\rho(\vec{r})$ as the density. Harriman [47] showed that any positive semi-definite $\rho(\vec{r})$ (i.e. $\rho(\vec{r}) \geq 0 \;\forall\; \vec{r} \in \mathbb{R}^3$) corresponds to *at least one* $|\Psi_\rho\rangle$, indicating that $\mathcal{Q}[\rho]$ is universally defined over all such

$\rho(\vec{r})$. Therefore, $E[\rho]$ is:

$$E[\rho] = \mathcal{Q}[\rho] - \int \rho(\vec{r}) v_{ext}(\vec{r}) d\vec{r} \tag{1.50}$$

over all positive semidefinite $\rho(\vec{r})$. In practice, constrained search over the exponentially scaling Hilbert space of many-electron wavefunctions $|\Psi_\rho\rangle$ is an impractical proposition. Nonetheless, this formalism provides a firmer formal justification for DFT, without considerations about ground state degeneracy or v-representability.

1.2.3 Kohn-Sham Theory

The formal knowledge that there exists an exact, variationally minimizable functional $E[\rho]$ is however of little practical benefit by itself, as seen from the failure of TFD theory. Somewhat paradoxically, the next major advance in DFT arose from **not** treating the kinetic energy T as an explicit functional of ρ. HF is capable of modeling T reasonably well, and so in 1965, Kohn and Sham (KS) proposed a scheme wherein the vast majority of T was approximated via a HF like expression through a Slater determinant intermediary [48].

Specifically, KS proposed looking at a fictitious system of noninteracting electrons (subjected to some local one-body potential $v_{KS}(\vec{r})$) that has the same $\rho(\vec{r})$ as the actual, interacting system of electrons. The fictitious noninteracting electron system can be easily solved, as the Hamiltonian $\mathbf{H}_{KS} = \mathbf{T}_e + \mathbf{V}_{KS}$ does not contain electron-electron repulsion, and can thus be exactly solved in a HF like manner. The ground eigenstate of $\mathbf{H}_{KS}$ is a single Slater determinant $\left|\Phi^{KS}\right\rangle$ with constituent occupied orbitals $\phi_i^{KS}, \phi_j^{KS} \ldots$ The resulting density is:

$$\rho(\vec{r}) = \sum_i |\phi_i(\vec{r})|^2 \tag{1.51}$$

The associated kinetic energy T_s for this non-interacting system is exactly:

$$T_s = \left\langle \Phi^{KS} \right| \mathbf{T_e} \left| \Phi^{KS} \right\rangle \tag{1.52}$$

KS proposed that:

$$E = T_s + E_{ne} + E_J + E_{xc} + V_{nn} \tag{1.53}$$

$$E_{xc} = E_x + E_c + T - T_s \tag{1.54}$$

where E_{ne} and E_J are explicit functionals of ρ given by Eqns 1.44 and 1.45, and E_{xc} is a catch-all term that includes exchange (E_x), correlation (E_c) and any difference in kinetic energy between the true interacting system (T) and the noninteracting kinetic energy (T_s). The last term should be small as HF is quite effective at modeling T, and T_s is computed similarly using a single Slater determinant. We note that T_s remains a functional of $\rho(\vec{r})$: just an implicit one as opposed to Eqn 1.46. The connection can be seen by noting that

$\rho(\vec{r})$ also completely specifies the solution to $\mathbf{H}_{KS}$ via the first HK theorem (since it does not need to assume electrons are interacting), and thus can be directly mapped to T_s.

In practice therefore, given some exchange-correlation functional E_{xc}, the Kohn-Sham procedure entails minimization of Eq 1.53 over some set of orbitals $\{\phi_i^{KS}\}$, very similar to the HF procedure. Indeed, the only difference is that the exchange term $\frac{1}{2}\sum\limits_{ij} K_{ij}$ is replaced by E_{xc}. KS procedures therefore have no higher a computational complexity than HF (often lower as K_{ij} need not be evaluated in local functionals, as noted in Sec 1.2.4).

KS theory therefore requires the following: the existence of one Slater determinant that has the same $\rho(\vec{r})$ as the actual system, and an effective E_{xc} functional. We will briefly examine both aspects here to understand the limitations of the method.

In a way, KS theory sidesteps the N-representability problem in that $\rho(\vec{r})$ arising from a single Slater determinant wavefunction $\left|\Phi^{KS}\right\rangle$ is guaranteed to be N-representable. The question therefore is whether *any* ground state $\rho(\vec{r})$ can be mapped to such $\left|\Phi^{KS}\right\rangle$. The answer intuitively appears affirmative, as the HF determinant can qualitatively approximate the true wavefunction and thus the density, even though it is not optimized for the latter. Therefore, it seems likely that it should be possible to construct a Slater determinant out of an arbitrary density (potentially through perturbing the HF orbitals). However, the existence of such a construction has not been exactly proven in practice. It has only been shown [47] that any arbitrary positive semidefinite $\rho(\vec{r})$ can be expressed as:

$$\rho(\vec{r}) = \sum_i f_i \left|\phi_i(\vec{r})\right|^2 \tag{1.55}$$

$$0 \leq f_i \leq 1 \tag{1.56}$$

A single Slater determinant construction requires that these occupation numbers f_i be either 0 or 1, as opposed to fractional values. The possibility of fractional f_i led to the development of models where $\rho(\vec{r})$ is reproduced by an ensemble of Slater determinants instead [49], although we will not consider such approaches further in this work.

Empirically, it has been shown the (near-)exact electron densities (from FCI or reasonable wavefunction approximations) can be reproduced by single Slater determinants for many chemical systems. However, exceptions are known for strongly correlated systems, where no one single Slater determinant can reproduce the exact wavefunction density [49]. Furthermore, empirical comparisons can only be done through some finite truncation of the Hilbert space (through use of grids or the so called ‘basis sets’), and the resulting observations need not be an accurate representation of the full Hilbert space behavior [50]. Realistically however, these comparisons are useful as practical calculations must utilize some Hilbert space truncation to avoid having to extend to the literal edge of the universe. From a pragmatic perspective therefore, it seems reasonable to conclude that the true densities of single reference chemical systems can be (reasonably well) represented by a single KS determinant, while the situation for multireference systems is potentially more complex.

The second issue with KS theory is whether E_{xc} can be sufficiently well approximated, as the exact functional is unknown and likely unknowable. The answer seemed to be no in 1965, as only exchange for UEG (given by Eqn 1.47) was known, and even the UEG correlation component remained intractable in the strong correlation limit. This led KS to pessimistically observe that they did "not expect an accurate description of chemical binding". Subsequent events would prove them wrong on this count, as KS-DFT would emerge as the most widely used quantum chemistry approximation, offering an unprecedented balance between predictive accuracy and low computational cost. One suspects Walter Kohn was probably not too unhappy about being proven incorrect, as he received one half of the 1998 Nobel prize in chemistry for his work on DFT [51].

1.2.4 Local Functionals

At present, density functional approximations (DFAs) for E_{xc} can broadly be classified into local and nonlocal, with the former being computationally simpler. For local functionals, E_{xc} is often expressed as:

$$E_{xc} = \int \rho(\vec{r})\epsilon_{xc} d\vec{r} \tag{1.57}$$

where ϵ_{xc} is the exchange-correlation energy density function. This is a local function of the density, in that the value of ϵ_{xc} for any given point $\vec{r}$ in space only depends on the value of ρ (and its derivatives) at that point only. In other words, ϵ_{xc} would not combine information from distant points in space together (unlike J_{ij} or K_{ij} whose integrands are functions of two electron positions $\vec{r}_1$ and $\vec{r}_2$). The locality of this approach makes local DFAs far more computationally affordable than HF, as the latter involves evaluation of nonlocal K integrals. However, local DFAs are often not very accurate, especially for evaluating nonlocal interactions like dispersion.

1.2.4.1 Local Density Approximation

The simplest of local functionals is the local density approximation (LDA), which assumes that E_{xc} can be just found from the local density $\rho(\vec{r})$ at any given point $\vec{r}$. Specifically:

$$E_{xc} = \int \rho(\vec{r})\epsilon_{xc}(\rho) d\vec{r} \tag{1.58}$$

where $\epsilon_{xc}(\rho)$ at any given point $\vec{r}$ is a function of only the value of the density $\rho(\vec{r})$ at that point. An example is the exchange functional given in Eqn 1.47, wherein $\epsilon_x(\rho) = -\frac{3}{4}\left(\frac{3\rho}{\pi}\right)^{\frac{1}{3}}$.

LDA however avoids any reference to spin. More useful is the local spin density approximation (LSDA), which utilizes the up and down spin densities ρ_α and ρ_β (with $\rho(\vec{r}) =$

$\rho_\alpha(\vec{r}) + \rho_\beta(\vec{r}))$. Since exchange only occurs between spins of the same sign, we can generalize Eqn 1.46 to:

$$E_x = \int \rho_\alpha(\vec{r})\epsilon_x^\alpha(\rho_\alpha)d\vec{r} + \int \rho_\beta(\vec{r})\epsilon_x^\beta(\rho_\beta)d\vec{r} \tag{1.59}$$

$$\epsilon_x^\alpha(\rho) = \epsilon_x^\beta(\rho) = -\frac{3}{2}\left(\frac{3\rho}{4\pi}\right)^{\frac{1}{3}} \tag{1.60}$$

Eq 1.59 is called Slater exchange, and using this minimal form for the full E_{xc} in KS theory yields significantly improved results relative to TFD. However, this model lacks correlation and is therefore nowhere near chemical accuracy.

The general LSDA model can be made exact for UEG if correlation is included. Since correlation exists between electrons of the same spin, as well as electrons of the opposite spin, it is not possible to separate the functional into pure spin-blocks (unlike Eqn 1.59). However, it is possible to have a general form wherein:

$$E_{xc} = \sum_{\sigma\in\{\alpha,\beta\}} \int \rho_\sigma(\vec{r})\epsilon_{xc}^\sigma(\rho_\alpha, \rho_\beta)d\vec{r} \tag{1.61}$$

where ϵ_{xc}^σ is associated with electrons of spin σ, which depends on both ρ_α and ρ_β.

Finding ϵ_{xc}^σ however entails knowing the correlation energy of the UEG for arbitrary $\rho_{\{\alpha,\beta\}}$. The correlation energy of UEG for large ρ (weakly correlated) limit could be estimated via the random phase approximation [52] (RPA) but the low ρ (strongly correlated) limit proved to be more elusive. In 1980, Ceperley and Alder provided diffusion Monte-Carlo (DMC) results for various ρ values of the UEG [53], paving the road for more accurate LSDA models. Soon after, Vosko, Wilk and Nusair presented the VWN models for LSDA correlation [54] via fitting to the results of Ref 53 and utilizing insight from RPA. In conjunction with Slater exchange, this yields SVWN functionals-which are perhaps the most popular LSDA models by far. It is nonetheless worth mentioning that Perdew and Wang presented an alternative LSDA parameterization in 1992 (named PW92), which has very similar performance to VWN but has emerged as the foundation for some of the most accurate modern functionals for ground state relative energy calculations [55]. Quantitatively speaking, SPW92 has an root mean square error (RMSE) of 66 kcal/mol for the total atomization energy (TAE) of non-multireference molecules in the highly accurate W4-11 dataset [56] (henceforth referred to as the TAE140nonMR dataset), while SVWN fares similarly [55]. This seems quite poor in light of the strength of a C-C bond ($\sim$ 100 kcal/mol) until we realize that HF has an RMSE of 125 kcal/mol, double that of LSDA [55]. The KS formalism thus permits a functional that is only exact for the UEG to significantly improve upon HF for computing chemically relevant ground state energy differences.

1.2.4.2 Generalized Gradient Approximation

The simplest useful generalization of the LSDA approach is to have ϵ_{xc} not only depend on the density, but also the density gradient. In other words:

$$E_{xc} = \sum_{\sigma\in\{\alpha,\beta\}} \int \rho_\sigma(\vec{r})\epsilon^\sigma_{xc}(\rho_\alpha, \rho_\beta, \vec{\nabla}\rho_\alpha, \vec{\nabla}\rho_\beta)d\vec{r} \tag{1.62}$$

, which is called the generalized gradient approximation (GGA). Inclusion of $\vec{\nabla}\rho_{\{\alpha,\beta\}}$ permits GGA functionals to look beyond a specific point $\vec{r}$, albeit only in the immediate neighborhood. Unlike the UEG, there is no exact model system for GGAs to be fit to, only exact constraints [57] and empirical data. GGA functionals have been consequently developed via empirical fitting, satisfaction of exact constraints or some intermediate strategy involving elements of both. Some popular GGAs are PBE [57] (non-empirical), BLYP [58] (empirical) and B97-D3 [59] (empirical). GGAs significantly improve upon accuracy relative to LSDA, with the three aforementioned functionals having RMSEs of 16 kcal/mol, 9 kcal/mol and 4.5 kcal/mol respectively [55], over the TAE140nonMR set. This is essentially an order of magnitude improvement over HF and LSDA.

1.2.4.3 Meta-GGAs

An even more general form of ϵ_{xc} can be obtained through inclusion of second derivative information, yielding:

$$Exc = \sum_{\sigma\in\{\alpha,\beta\}} \int \rho_\sigma(\vec{r})\epsilon^\sigma_{xc}(\rho_\alpha, \rho_\beta, \vec{\nabla}\rho_\alpha, \vec{\nabla}\rho_\beta, \nabla^2\rho_\alpha, \nabla^2\rho_\beta)d\vec{r} \tag{1.63}$$

which is called a meta-GGA (mGGA). For KS-DFT, there is another second derivative based quantity called τ, given by:

$$\tau(\vec{r}) = \sum_i \left|\vec{\nabla}\phi_i(\vec{r})\right|^2 \tag{1.64}$$

This is called the kinetic energy density as it is positive semidefinite and:

$$T_s = -\frac{1}{2}\sum_i \int \phi_i^*(\vec{r})\nabla^2\phi_i(\vec{r}) = \frac{1}{2}\sum_i \int \left|\vec{\nabla}\phi_i(\vec{r})\right|^2 d\vec{r} = \frac{1}{2}\int \tau(\vec{r})d\vec{r} \tag{1.65}$$

via integration by parts (and because the orbitals and their derivatives are all zero at infinity). τ is related to $\nabla^2\rho$ via [60]:

$$\tau = \frac{\nabla^2\rho}{2} - \sum_i \phi_i^*(\vec{r})\nabla^2\phi_i(\vec{r}) \tag{1.66}$$

and so only one of τ or $\nabla^2\rho$ is used in mGGAs. In practice, τ is more popular, and most modern mGGAs utilize it over $\nabla^2\rho$. So a modern mGGA functional is typically of the form:

$$Exc = \sum_{\sigma\in\{\alpha,\beta\}} \int \rho_\sigma(\vec{r})\epsilon^\sigma_{xc}(\rho_\alpha, \rho_\beta, \vec{\nabla}\rho_\alpha, \vec{\nabla}\rho_\beta, \tau_\alpha, \tau_\beta)d\vec{r} \tag{1.67}$$

It should be noted that while the purely ρ dependent terms are constrained to LSDA, there are not enough exactly known constraints to anchor all of the $\vec{\nabla}\rho$ and τ dependent terms [61, 62]. The space of all possible mGGAs is thus enormous (exponentially larger than GGAs) [63], and it is critical to avoid overfitting when functionals are designed empirically. Conversely, absence of specific datatypes in the training set (such as thermochemistry or barrier heights) may make it difficult for the trained DFA to be transferable to new kinds of data, as it might be stuck to a local minimum in the very high dimensional space of possible functionals, that is quite far from the exact functional.

A popular modern meta-GGA is the Strongly Constrained and Appropriately Normed (SCAN) functional [62], which satisfies all the exact constraints known for mGGAs and was also fit to atomic energies. This has an RMSE of 3.6 kcal/mol over the TAE140nonMR set [55], which is only a minor improvement over the empirical B97-D3 GGA, but a significant improvement relative to the nonempirical PBE GGA. Another successful mGGA is the semi-empirically designed B97M-V functional [63], which was devised through combinatorially finding the most transferable form in mGGA space. This has an RMSE of 3.9 kcal/mol for TAE140nonMR [55], indicating a small but steady improvement over GGAs for this dataset.

Local functional design seldom adds third or higher order derivatives to ϵ_{xc}. The reason for this is twofold. The mGGA space is already too vast to easily explore, and adding further variables would only complicate matters further. Secondly, local functionals infamously suffer from a systematic problem called delocalization error that is unlikely to be solved by adding higher order local derivatives. Consequently, a new recipe for functional development is needed. Before proceeding along that direction however, it feels necessary to discuss delocalization error briefly, and how it might be remedied.

1.2.5 Delocalization Error[1]

Perdew et al. [64] showed that the electronic energy of a system with fractional charge is exactly determined by a linear interpolation between the energies corresponding to the two closest integer electron numbers. Mathematically, therefore, the energy E of a system with $m-x$ electrons (where m is a nonnegative integer and x lies between 0 and 1) is given by:

$$E(m-x) = E(m) + x(E(m-1) - E(m)) \tag{1.68}$$

---

[1] The content and figures of this subsection have been adapted with permission from **Hait, D.**; Head-Gordon, M. "Delocalization errors in density functional theory are essentially quadratic in fractional electron number." *J. Phys. Chem. Lett.*, **9**, 6280-6288. 2018. Copyright 2018 American Chemical Society.

Eqn. 1.68 therefore specifies that the electronic energy is piecewise linear with respect to the electron number. It does not specify the energies for integer electron numbers themselves, but difference between molecular electron affinity (EA), $E(m+1) - E(m)$, and ionization potential (IP) $E(m-1) - E(m)$, means that a derivative discontinuity in the energy as a function of electron number occurs at m [64]. This is illustrated by the exact curve for the F atom, on the left-hand panel of Fig 1.1.

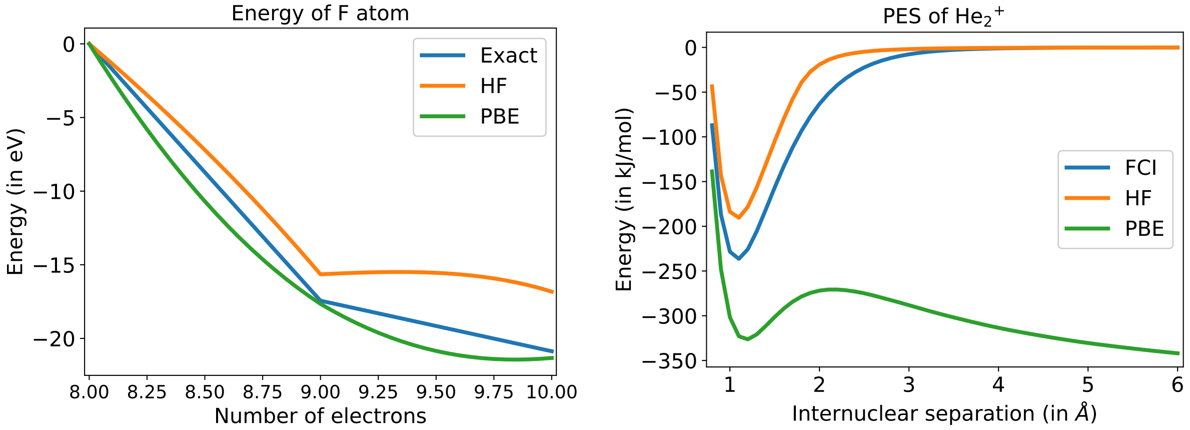


Figure 1.1: Effect of delocalization error on fractional charges (left) and dissociation of charged species (right), showing overestimation of energy by HF and underestimation by a typical DFA (PBE). The reported energies are relative to the F^+ cation on the left panel and the true dissociation limit of the He_2^+ complex (He^++He) on the right panel, as predicted by each method. ‘Exact’ behavior on the left panel is obtained from linear interpolation between CCSD(T) energies. The aug-pc-4 basis [65–69] was employed for HF and PBE on the left panel, while all calculations on the right panel used the aug-cc-pVTZ basis [70].

Several other proofs of piecewise linearity have been given elsewhere [71–73]. A particularly simple proof is that of Yang et.al. [71] who constructed a supersystem of q identical non-interacting copies of an m-electron molecule, from which $p < q$ electrons are removed. From one perspective the supersystem is therefore p copies of $E(m-1)$ and $q-p$ copies of $E(m)$. Equivalently, as mixing degenerate states does not affect the energy, the supersystem can be composed of q identical fragments, each with a fractional electron $x = p/q$ removed, with total energy that is evidently:

$$qE\left(m-x\right) = \left(q-p\right)E\left(m\right) + pE\left(m-1\right) \tag{1.69}$$

Dividing through by q yields Eqn. 1.68 for rational x. This was generalized to irrational x by Ayers [72].

Most approximate functionals fail to satisfy Eqn 1.68, giving energies for fractional electron numbers that are typically too low [74, 75] (as can be seen on the left panel of Fig 1.1). This error is often explained in terms of electron self-interaction. E_J contains a spurious self-repulsion energy (as evident from Eqn 1.45) which E_{xc} in principle should cancel. The correction however is very approximate for most functionals, on account of the locality of E_{xc}, while self-repulsion is explicitly nonlocal. The electron density subsequently delocalizes (becomes less compact) in order to minimize the residual self-repulsion, and this tends to generate fractional charges whose energies are lower than what Eqn 1.68 predicts. This erroneous delocalization behavior leads to catastrophes like incorrect asymptotic behavior in dissociation curves for charged species [76] (as shown in the right panel of Fig 1.1), fractional electrons dissociating from anions [77] or spurious fractional charges at the dissociation limit for polar bonds [78–81]. It also causes other complications like lowering of barrier heights [82–84] and reduction of band gaps [85, 86].

Self-repulsion alone is however an inadequate explanation for deviation from Eqn 1.68 [83, 87]. As noted before, HF lacks electron self-repulsion, but nonetheless does not adhere to Eqn 1.68 aside from the trivial single electron cases. In fact, HF predicts *too high* an energy for fractional charges on account of missing correlation energy, as proven in Ref 88 and also empirically evident from the concave HF curves on the left panel of Fig 1.1. This effect is often termed as ‘many electron self-interaction error’ [83], though it is somewhat misleading as electrons are not really interacting with themselves in this case! We will therefore use the term ‘self-interaction error’ to only refer to incorrect energy predictions arising from an electron spuriously repelling itself, and reserve the more general term ‘delocalization error’ for any and all deviations from Eqn 1.68. Delocalization error in HF manifests itself through overlocalization of electron density, leading to incorrect behavior such as too high barrier heights [84] and too quick a decay in dipole moments during bond dissociation [80]. Effects of this can be seen on the left panel of Fig 1.1, where HF predicts too high an energy for systems with fractional charges due to missing correlation energy.

1.2.6 Hybrid Functionals

The preceding section reveals that HF tends to overlocalize densities while local DFAs overdelocalize them. One potential middle path is to hybridize [89] HF and local DFT to create a generalized KS approach. Specifically, given some local exchange correlation functional $E_{xc} = E_x + E_c$, we can develop a global hybrid functional:

$$E_{xc} = -\frac{c}{2}\sum_{ij} K_{ij} + (1-c)E_x[\rho] + E_c[\rho] \tag{1.70}$$

In other words, a portion of HF exchange is added in, and a corresponding amount of the local exchange E_x is removed (to prevent double-counting of exchange and to recover the UEG limit). This reduces delocalization error via reduction of the self-repulsion component in E_J through HF exchange. We note that use of pure HF exchange through $c = 1$ generally leads to suboptimal results with local E_c [55], despite the formal lack of self-repulsion.

Global hybrid functionals of the form given by Eqn 1.70 are arguably the most widely used DFAs in quantum chemistry [90]. The two most popular ones are B3LYP [89] and PBE0 [91], derived from the local GGAs BLYP and PBE respectively. PBE0 in particular has quite a simple form, being composed of 25% HF exchange, 75% PBE exchange and 100% PBE correlation (i.e. $c = 0.25$). Both B3LYP and PBE0 have RMSE of ~ 4.5 kcal/mol over the TAE140nonMR dataset [55], which is a significant improvement over the 9 and 16 kcal/mol RMSE of the parent GGAs, to say nothing of the 125 kcal/mol RMSE of HF! We note that typical values of c range from $0.1 - 0.4$, though larger values may be relevant for specific applications like barrier height calculations where delocalization error is the major issue [92].

Delocalization error can be further reduced via use of a range-separated hybrid [93, 94] (RSH) scheme instead of a global hybrid approach. RSH models adds a distance dependent amount of HF exchange to ϵ_{xc}, instead of the fixed amount set by c in Eqn 1.70. Specifically, RSH functionals usually utilize K'_{ij} defined by:

$$K'_{ij} = \int \phi_i^*(\vec{r}_1, s_1)\phi_j^*(\vec{r}_2, s_2)\frac{c + (1-c)\mathrm{erf}\left(\omega\left|\vec{r}_1 - \vec{r}_2\right|\right)}{\left|\vec{r}_1 - \vec{r}_2\right|}\phi_j(\vec{r}_1, s_1)\phi_i(\vec{r}_2, s_2)d\vec{r}_1 ds_1 d\vec{r}_2 ds_2 \tag{1.71}$$

The modified interelectron repulsion operator is $\dfrac{c + (1-c)\mathrm{erf}\left(\omega\left|\vec{r}_1 - \vec{r}_2\right|\right)}{\left|\vec{r}_1 - \vec{r}_2\right|}$ which is merely $c\dfrac{1}{\left|\vec{r}_1 - \vec{r}_2\right|}$ when $\left|\vec{r}_1 - \vec{r}_2\right| \to 0$ and $\dfrac{1}{\left|\vec{r}_1 - \vec{r}_2\right|}$ when $\left|\vec{r}_1 - \vec{r}_2\right| \to \infty$. The DFA is then

$$E_{xc} = -\frac{1}{2}\sum_{ij} K'_{ij} + E'_x[\rho] + E_c[\rho] \tag{1.72}$$

where E'_x is an adjusted exchange functional dependent on c, ω that should prevent double counting of exchange and help recover the UEG limit. RSH DFAs have only a fraction c of the HF exchange present at short interelectron separations, but that fraction goes to 100% HF at the infinite separation limit. The parameter ω controls the rate of the transition and is generally $0.2-0.5$ a.u. The c parameter is generally slightly smaller than corresponding global hybrids, normally ranging between $0.1-0.3$. Commonly used RSH functionals almost always have lower delocalization error than common global hybrids (see Ref 95 for comparisons), and are thus increasingly popular. As an example, the modern ωB97M-V functional [96] (designed by combinatorial functional design like B97M-V) has an RMSE of 2.2 kcal/mol over the TAE140nonMR dataset [55], almost attaining chemical accuracy.

One important point about hybrid functionals is that they have the same computational scaling as HF due to the need to evaluate K_{ij} or related quantities like K'_{ij}. Their formal $O(N^4)$ scaling is thus a significant increase in computational complexity from the $O(N^3)$ cost of local functionals. However, the resulting increase in accuracy makes this worthwhile for molecular systems, to the extent that hybrid DFAs have become the default choice in almost all circumstances. For extended solids however, the greater computational complexity leads to limited use of hybrid functionals (despite recognition of their higher accuracy) and local functionals remain the usual protocol.

1.2.7 Double Hybrids

The hybridization of HF and local DFAs leading to hybrid functionals inspired the development of double hybrid functionals that attempt to combine local DFAs with correlated wavefunction methods [97]. The most common route is to add MP2 correlation, with the general form of a global double hybrid being the following (in analogy to Eqn. 1.70):

$$E_{xc} = -\frac{c_1}{2}\sum_{ij} K_{ij} + (1-c_1)E_x[\rho] + c_2 E^{(2)} + (1-c_2)E_c[\rho] \tag{1.73}$$

where $E^{(2)}$ is the MP2 correlation energy, and c_1, c_2 are parameters that control mixing between local and nonlocal elements. A similar form can be defined for RSH double hybrid functionals. It is also possible to use correlation energy from other methods like RPA or CCSD to create double hybrids, but we do not explore such DFAs in this work.

One interesting property of double hybrid functionals is that c_1 is generally larger than what is normally used for hybrid functionals, being generally in the $0.6 < c_1 < 0.9$ range. The use of a high fraction of HF exchange in conjunction with a significant amount of nonlocal correlation permits double hybrid functionals to have rather low levels of delocalization error. We refer the interested reader to Ref 81 for a quick overview about various ways to define double hybrid functionals and obtain orbitals for evaluation of $E^{(2)}$.

Evaluation of $E^{(2)}$ is of $O(N^5)$ cost, making double hybrid functionals even more computationally demanding than hybrid functionals or HF. Nonetheless, there is a significant improvement in accuracy, with the TAE140nonMR RMSE of the combinatorially designed ωB97M(2) functional being 1.6 kcal/mol [98], which takes it halfway to chemical accuracy relative to the ωB97M-V functional. For context, we note that MP2 has an RMSE of 9.6 kcal/mol over the same set [98] and CCSD has 10.8 kcal/mol [56], comparable to BLYP and worse than mGGAs or hybrid functionals discussed so far. Indeed, double hybrid functionals are some of the most accurate density functional approximations that have yet been developed, and offer the closest glimpse of chemical accuracy from a DFT standpoint. Their use is not as routine as hybrid DFAs, but advances in computing MP2 correlation more efficiently [19] are likely to enhance their practical utility in the days to come.

1.2.8 Jacob's Ladder

At present there are a few hundred DFAs reported in the literature [55], forming a veritable alphabet soup that is difficult to navigate even for experts. For ease of classification, the community has by and large agreed to define a hierarchy of DFAs based on the complexity of E_{xc}. This is generally called Jacob's ladder (based on a suggestion from Perdew [99]), which extends from the 'hell' of no E_{xc} to the 'heaven' of chemical accuracy. Fig 1.2 provides a visual representation of the ladder, and much of the jargon therein has been already explained earlier. To recapitulate briefly, the complexity of the functional form increases upon climbing the ladder. It is hoped that the accuracy also increases upon ascent, although in practice

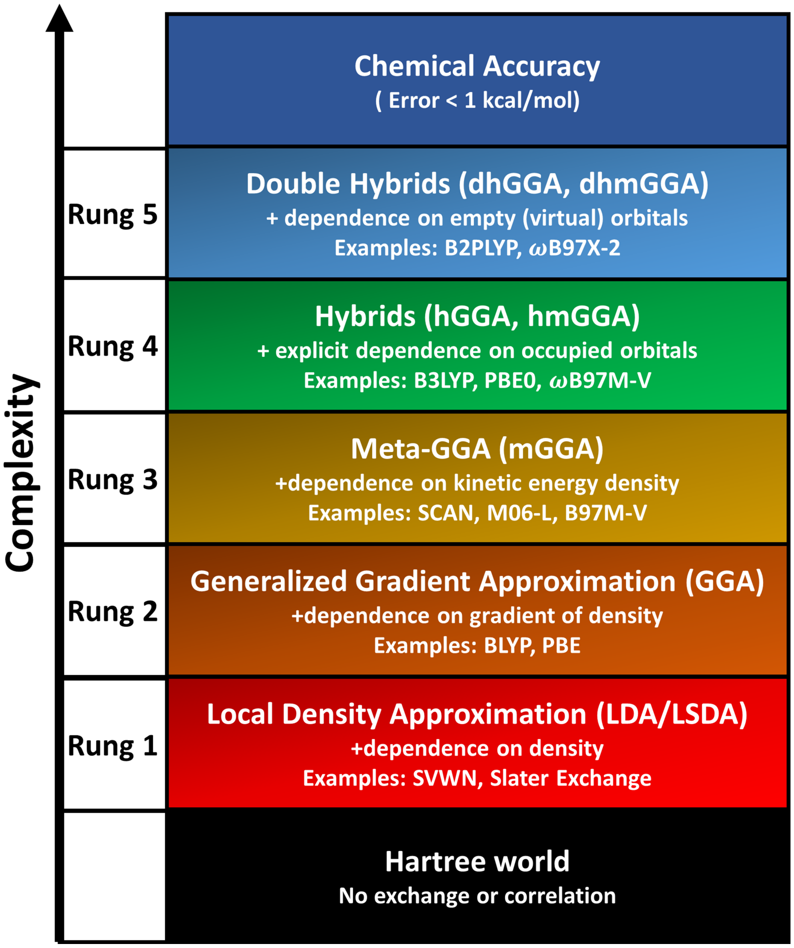

Chemical Accuracy
(Error < 1 kcal/mol)
Rung 5
Double Hybrids (dhGGA, dhmGGA)
+ dependence on empty (virtual) orbitals
Examples: B2PLYP, ωB97X-2
Rung 4
Hybrids (hGGA, hmGGA)
+ explicit dependence on occupied orbitals
Examples: B3LYP, PBE0, ωB97M-V
Rung 3
Meta-GGA (mGGA)
+dependence on kinetic energy density
Examples: SCAN, M06-L, B97M-V
Rung 2
Generalized Gradient Approximation (GGA)
+dependence on gradient of density
Examples: BLYP, PBE
Rung 1
Local Density Approximation (LDA/LSDA)
+dependence on density
Examples: SVWN, Slater Exchange
Hartree world
No exchange or correlation
Complexity

Figure 1.2: Jacob's ladder for DFAs.

the best one could hope for is to have the best functional in any rung be more accurate than the best functionals in all rungs below.

LSDA constitutes rung 1 as it is the simplest functional form (only depending on local spin density) and is exact only for UEG. GGAs form rung 2 as they include information about the local spin density gradients as well, while mGGAs form rung 3 due to the inclusion of local spin kinetic energy densities in the functional (or other second derivative information). The computational cost of E_{xc} increases on ascending from rungs 1 to 3, but the asymptotic scaling is unaffected as it is dominated by the need to diagonalize a Fock matrix (of size $O(N)$) over many self-consistent iterations. The local functionals therefore all have an $O(N^3)$ asymptotic cost, which is lower than HF, and has permitted widespread use in condensed phase calculations. They are however plagued by delocalization error, and are usually not very accurate relative to correlated wavefunction methods (despite surpassing HF). That said, modern mGGAs like SCAN and B97M-V have proven to be more accurate for many chemical problems than what was historically expected of local functionals, even approaching hybrid functional levels of accuracy at times [55].

Rung 4 is made of hybrid functionals, wherein HF exchange energy (or a RSH equivalent) is computed from occupied KS orbitals directly, without explicit recourse to the density. Such approaches are a mixture of local DFAs from rungs 1-3 and HF, but surpass either in accuracy. Indeed, Rung 4 functionals are generally quite a bit more accurate than MP2, and are even comparable to CCSD at times for relative energies. These functionals however have $O(N^4)$ computational complexity, due to the need to evaluate nonlocal HF exchange.

Rung 5 represents the top level of Jacob's ladder as accessed today, and consists of functionals which use the unoccupied orbitals to compute a nonlocal, wavefunction based contribution to the correlation energy. This estimate is generally obtained from MP2, leading to $O(N^5)$ scaling computational cost that exceeds HF. However the resulting increase in accuracy is significant, with several rung 5 methods statistically surpassing CCSD for ground state relative energy calculations.

1.2.9 Dispersion Interactions

A perceptive reader might have noticed that an entire class of interactions were neglected in the discussion of Jacob's ladder. Dispersion interactions are increasingly being recognized as key players between large, polarizable systems in nature, such as possibly playing a role in the clustering of hydrocarbons to form soot [100]. Dispersion involves interactions between induced dipoles in nonpolar systems, and a net attractive force is only possible if the induced dipoles are correlated with each other over long distances. The local density models in rungs 1-3 are thus incapable of describing long range dispersion, although some suggestions have been made that some highly parameterized modern functionals might mimic intermediate distance effects [101] where there is still perceptible overlap between electron clouds of fragments [102]. Nonetheless, the long-range R_{AB}^{-6} scaling of dispersion interaction energy (vs interspecies distance R_{AB}) is not reproduced. Hybrid functionals confer no advantage, as HF lacks correlation. Double hybrids do possess nonlocal correlation and are thus more useful for such applications. However, Eqn 1.73 scales down nonlocal correlation and a considerable amount of dispersion is therefore typically missed. Steps in a direction orthogonal to Jacob's ladder might consequently be necessary for describing dispersion.

There are two broad classes of solutions to this problem. The first was proposed by Grimme [103], wherein explicit, molecular mechanics like $-\frac{C_6}{R_{AB}^6} f_{damp}(R_{AB})$ terms were added to the system energy. The damping function $f_{damp}(R_{AB})$ prevents double counting of correlation at small R_{AB}, where electron clouds start to overlap. This DFT-D approach was subsequently augmented through inclusion of other terms and dependence on coordination numbers to create the DFT-D3 approach [104], which is perhaps the most widely used class of dispersion corrections. These corrections are classical and semi-empirical, but are computationally simple and widely applicable. It is possible to fit D3 short-range damping parameters during functional training as well, permitting quite good results in many cases. Indeed, -D3 corrections are available for virtually any available functional. For example, B3LYP-D3 has an RMSE of 0.4 kcal/mol vs 3.8 kcal/mol for bare B3LYP, for a set of 66 binding energies of non-covalent interactions [55] found in organic molecules and biomolecules (S66 [105]). Note that these interactions are weak enough that the thermochemistry inspired definition of chemical accuracy as 1 kcal/mol is inadequate, and methods should instead aim to have intermolecular binding energies accurate to 0.1 kcal/mol instead. A DFT-D4 scheme has also been proposed very recently [106], which aims to use semiempirical QM charges in the dispersion calculation instead of the purely classical approach earlier.

The second approach is explicitly density dependent, wherein a nonlocal correlation term is defined by $\int \rho(\vec{r}_1)\Phi(\vec{r}_1,\vec{r}_2)\rho(\vec{r}_2)d\vec{r}_1 d\vec{r}_2$. The nonlocal kernel $\Phi(\vec{r}_1,\vec{r}_2)$ therefore correlates densities at $\vec{r}_{\{1,2\}}$, permitting dispersion interactions. Perhaps the most successful of such nonlocal correlation (NLC) functionals is the VV10 functional [107] by Vydrov and Van Voorhis. The training of the modern ωB97X-V [108], B97M-V and ωB97M-V functionals included VV10 (denoted by the V at the end), which is largely responsible for their accuracy in modeling noncovalent interactions. Indeed, the three functionals have RMSEs of 0.13 kcal/mol, 0.18 kcal/mol and 0.15 kcal/mol for the S66 dataset. Even the ωB97M(2) double hybrid utilizes a fraction of VV10 to augment MP2 for dispersion interactions, and fares comparably to ωB97M-V. It is nonetheless worth noting that evaluation of NLC correlation via the double integral is computationally demanding, which is perhaps why NLC functionals are not quite routinely available in quantum chemistry packages.

1.2.10 Spin Symmetry Breaking

A chemical system with equal numbers of up and down spin electrons should have:

$$\rho_\alpha(\vec{r}) = \rho_\beta(\vec{r}) = \frac{1}{2}\rho(\vec{r}) \tag{1.74}$$

in the absence of magnetic fields, or other spin-dependent terms in the Hamiltonian (as there is no bias towards any particular direction of spin). The exact functional should also reproduce this behavior. DFAs however sometimes minimize E for such systems with spin symmetry broken determinants, where $\rho_\alpha(\vec{r}) \neq \rho_\beta(\vec{r})$ for all $\vec{r} \in \mathbb{R}^3$. This behavior is most common in multireference systems, as existing DFAs are generally not very effective at modeling strong electron correlation and therefore face difficulty in modeling the system with a single, spin unpolarized Slater determinant. This is similar to HF also obtaining spin polarized determinants [14] as the lowest energy solution for such systems, arising from an attempt to unpair strongly interacting electron pairs within a single determinant framework.

For the ground state energies of systems with spin symmetry breaking, it is typical to utilize the lower energy arising from the spin symmetry broken unrestricted KS (UKS) solution over a higher energy obtained from a symmetry preserving spin-restricted KS (RKS) solution. This is not necessarily a problem as the KS solution need only have correct $\rho(\vec{r})$, without any restrictions on $\rho_\alpha(\vec{r}), \rho_\beta(\vec{r})$. Nonetheless, the presence of spin symmetry breaking is a good sign for caution when working with most existing DFAs. This behavior has been connected to ‘fractional spin’ errors in existing DFAs [109], analogous to the ‘fractional charge’/delocalization error described earlier. Chapters 7 and 10 discuss some challenges that arise from working with spin symmetry broken UKS solutions. However, the variational principle from the HK theorems, and experience with spin symmetry breaking in HF [14, 15] indicates that use of the lowest energy UKS solutions for ground state energies is generally the optimal option.

1.2.11 Beyond Ground State Energies

It is universally acknowledged today that DFT has come a long way since the KS formalism was first presented in 1965. Initial scepticism about applicability to chemistry has been brushed aside by decades of successful use [55], to the point that KS-DFT is the default electronic structure paradigm for most applications [110]. While chemical accuracy has not yet been attained, a combination of physically motivated functional design and use of vast quantities of high level wavefunction theory data has resulted in accurate functionals that are very useful for making chemical predictions, and come closer to the heaven atop Jacob’s ladder than ever before [55, 98, 111–113]. Data driven functional design [98] and machine learning approaches [114] will continue to churn out DFAs with ever increasing accuracy to join the functional soup.

At the same time, all of these developments to date have near exclusively focused on improving the prediction quality for ground state energy differences. The centrality of those to chemistry can hardly be denied, and there is virtually no reason to develop a functional that is suboptimal for predicting ground state energies (aside from some novelty value). However, there is certainly more to chemistry and the exact functional than the electronic ground state energy. The birth of QM owes a great deal to spectroscopy, and experimental spectra continue to be the greatest source of validation for the theory (within the domain of chemical physics). And yet, very little spectroscopy can be done with the knowledge of the ground state energy alone. Modeling microwave spectra requires accurate molecular geometries for rotational constants, predicting infrared spectra needs accurate vibrational frequencies, understanding ultraviolet/visible spectra benefits from accurate description of valence electron excitations, and X-ray spectra cannot be properly computed without accounting for core-hole relaxations. All of these techniques also require understanding the interaction between matter and the electric field of light. Nuclear magnetic resonance (NMR) is perhaps even trickier as it entails response to magnetic fields!

In the years past, experimental attention often remained focused on small molecules that were coincidentally tractable with high level wavefunction methods and not much DFT was necessary. However, there are rather few such molecules left, and laboratory apparatus have only gotten more sophisticated. It is now broadly acknowledged that the frontiers of chemical research lie in complex systems. A great deal of modern experiments now probe the condensed phase, where few practical QM options exist outside of DFT. Even in the molecular realm, studies are shifting to larger and larger species for which DFT is the only computationally tractable approach with even a ghost of a chance of success.

Even from a purely theoretical perspective, there is much about DFT that remains underexplored due to the strong existing focus on ground state relative energies. For instance, the accuracy of DFT densities has only recently received any attention [115, 116], even though ρ is as central to the theory as E! A functional that predicts reasonable E through inaccurate ρ is hardly acceptable in a formal sense. Accurate molecular properties that are defined by response to external fields are also of interest, as they reveal how transferable DFAs are to scenarios with different local potentials that are formally covered by the HK theorems and the KS approach. Finally, while the ground state density corresponds to the global minimum of the exact energy functional, it is also known that other extremal densities of that functional correspond to electronic excited states [117]. Understanding the contours of functionals around excited state extrema has the potential to reveal much that has not been previously explored.

It would therefore behoove us to go beyond the ground state energy and understand how existing DFAs fare at modeling field-matter interactions, through molecular properties and excited states. Such understanding would lay the groundwork for successful synergy of theory with experiment, and would also permit quantum chemists to resolve shortcomings of existing approximations. This is the principal objective that this dissertation hopes to contribute to. In particular, half of this dissertation is devoted to understanding DFT predictions for molecular properties based on electronic densities (dipole moments and second spatial cumulant of electron density) and the response of such properties to electric fields (static dipole polarizability). The other half is concerned with modeling electronic excited states with DFT, beginning with the characterization of the failure of widely used linear-response approaches for describing excited state bond dissociation. It subsequently focuses on the development of an orbital optimization protocol for excited states, which can be employed to compute X-ray spectra in excellent agreement with experiment (unlike the standard linear-response approach). The outline in the subsequent pages provide greater details about the composition of each half of the dissertation.

Chapter 2

Outline of the Dissertation

This dissertation examines the utility of DFT for computing density dependent properties, and electronic excited states. The first half of this dissertation assesses the ability of modern DFT to predict molecular dipole moments, second cumulants of electron density and static dipole polarizabilities, all of which are properties connected to the response of the molecular energy to external, time-independent electric fields. The second half of the dissertation examines DFT approaches to molecular excited states. In particular, it focuses on excited state DFT approaches that entail state specific orbital optimization. An efficient and reliable algorithm for excited state specific orbital optimization is presented. This approach is found to yield semiquantitative accuracy against experiment for molecular core-level excitations, indicating that it can be readily applied to interpret experimental time-resolved X-ray absorption studies of chemical dynamics, particularly for large systems.

This dissertation is organized as follows:

Chapter 3: Foreword to Electrical Response Properties

This chapter discusses the role of accurate densities in DFT, and how density predictions cannot be decoupled for energy predictions when external potentials are applied. Density information therefore can be an important class of new data used to develop and assess DFAs. A convenient category of density related data are molecular multipole moments. Multipole moments are the first-order responses of the energy to spatial derivatives of the electric field strength. Stronger field strengths significantly perturb the electron density as well, with the effects being described by multipole polarizabilities. The quality of density functional theory (DFT) prediction of molecular multipole moments and polarizabilities thus characterizes errors in modeling the electron density itself, as well as the performance in describing molecules interacting with external electric fields. Benchmark datasets of multipole moments and polarizabilities therefore can be useful for functional development, as well as for understanding where present methods fall short.

Chapter 4: Dipole Moments

The content and figures of this chapter are reprinted or adapted with permission from **Hait, D**.; Head-Gordon, M. "How accurate is density functional theory at predicting dipole moments? An assessment using a new database of 200 benchmark values." *J. Chem. Theory Comput.*, **14**, 1969-1981. 2018. Copyright 2018 American Chemical Society.

The dipole moment is the simplest nontrivial multipole moment, representing the first spatial moment of the electron density. This chapter assesses the the accuracy of modern density functionals for calculating dipole moments. A database of 200 benchmark dipole moments is developed using coupled cluster theory through triple excitations, extrapolated to the complete basis set (CBS) limit. This new database is used to assess the performance of 88 popular or recently developed density functionals. The results suggest that double hybrid functionals perform the best, yielding dipole moments within $\sim$ 3.6-4.5% regularized RMS error vs the reference values—which is not very different from the 4% regularized RMS error produced by CCSD. Many hybrid functionals also perform quite well, generating regularized RMS errors in the 5-6% range. Some functionals however exhibit large outliers and local functionals in general perform less well than hybrids or double hybrids.

Chapter 5: Second Cumulants of the Electron Density

The content and figures of this chapter are reprinted or adapted from **Hait, D**.*; Liang, Y.H.*; Head-Gordon, M. "Too big, too small or just right? A benchmark assessment of density functional theory for predicting the spatial extent of the electron density of small chemical systems." *J. Chem. Phys.,* **154**, 074109, 2021; with full permission from AIP Publishing and all co-authors. (* indicates equal contribution from authors.)

The quadrupole moment of polar molecules is unsuitable for direct investigation as only the lowest non-zero multipole moment of a molecule is translationally invariant, making the higher-order moments origin-dependent. This chapter therefore focuses on the translationally invariant 3×3 matrix of second *cumulants* (or spatial variances) of the electron density as the quantity of interest (denoted by $\boldsymbol{\mathcal{K}}$), instead of the 3×3 quadrupole moment matrix. The principal components of $\boldsymbol{\mathcal{K}}$ are the square of the spatial extent of the electron density along each axis. A benchmark dataset of the principal components of $\boldsymbol{\mathcal{K}}$ for 100 small molecules at the CCSD(T)/CBS level is developed, resulting in 213 independent $\boldsymbol{\mathcal{K}}$ components. The performance of 47 popular and recent density functionals is assessed against this Var213 dataset. Several functionals, especially double hybrids, and also SCAN and SCAN0 yield reliable second cumulants, although some modern, empirically parameterized functionals yield more disappointing performance. The H and Be atoms in particular are challenging for nearly all methods, indicating that future functional development could benefit from inclusion of their density information in training or testing protocols.

Chapter 6: Static Polarizabilities

The content and figures of this chapter are reprinted or adapted from **Hait, D.**; Head-Gordon, M. "How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry." *Phys. Chem. Chem. Phys.*, **20**, 19800-19810. 2018; with permission from the PCCP Owner Societies.

Static polarizabilities are the first response of the electron density to electric fields, and are therefore important for predicting intermolecular and molecule-field interactions. They also offer a global measure of the accuracy of the treatment of excited states by density functionals in a formally exact manner. This chapter presents a database of benchmark static polarizabilities for 132 small species at equilibrium geometry, using CCSD(T)/CBS, for the purpose of developing and assessing density functionals. The performance of 60 popular and recent functionals are also assessed, which indicates that double hybrid functionals perform the best, having RMS relative errors in the range of 2.5-3.8% . Many hybrid functionals also give quite reasonable estimates with 4-5% RMS relative error. A few meta-GGAs like mBEEF and MVS yield performance comparable to hybrids, indicating potential for improved excited state predictions relative to typical local functionals. Some recent functionals however are found to be prone to catastrophic failure (possibly as a consequence of overparameterization), indicating a need for caution in applying these.

Chapter 7: Polarizabilities Beyond the Coulson-Fischer Point

The content and figures of this chapter are reprinted or adapted from **Hait, D.***; Rettig, A.*; Head-Gordon, M. "Well-behaved versus ill-behaved density functionals for single bond dissociation: Separating success from disaster functional by functional for stretched H_2" *J. Chem. Phys.*, **150**, 094115. 2019; with full permission from AIP Publishing and all co-authors. (* indicates equal contribution from authors.)

Unrestricted DFT methods are typically expected to describe the homolytic dissociation of nonpolar single bonds in neutral species with qualitative accuracy, due to the lack of significant delocalization error. This chapter reports that many widely used density functional approximations fail to describe features along the dissociation curve of the simple H_2 molecule. This is not a universal failure of DFT in the sense that many classic functionals like PBE and B3LYP give very reasonable results, as do some more modern methods like MS2. However, some other widely used functionals like B97-D (empirically fitted) and TPSS (non-empirically constrained) predict qualitatively wrong static polarizabilities, force constants and some even introduce an artificial barrier against association of independent H atoms to form H_2. The polarizability and force constant prediction failures appear to stem from incomplete spin localization into individual H atoms beyond the Coulson-Fischer point, resulting in 'fractionally bonded' species where the ionic contributions to the Slater determinant are not completely eliminated, unlike the case of unrestricted Hartree-Fock. These errors therefore appear to be a consequence of poor self-consistent spin density prediction by the problematic functional. The same reasons could potentially lead to spurious barriers towards H atom association, indirectly also leading to incorrect forces. These unphysicalities suggest that the use of problematic functionals is probably unwise in *ab-initio* dynamics calculations, especially if strong electrostatic interactions are possible.

Chapter 8: Afterword on Electrical Response Properties

This chapter summarizes the results of the preceding four chapters. The successes and some potential shortcomings of the presented material is discussed. The future use of response properties in DFA development and other applications is touched upon.

Chapter 9: Foreword on Electronic Excited States

Some of the content and figures of this chapter are reprinted or adapted with permission from two works. They are:

1. **Hait, D**.; Head-Gordon, M. "Orbital Optimized Density Functional Theory for Electronic Excited States." *J. Phys. Chem. Lett.,* **12**, 4517-4529. 2021.
2. **Hait, D**.*; Rettig, A.*; Head-Gordon, M. "Beyond the Coulson-Fischer point: characterizing single excitation CI and TDDFT for excited states in single bond dissociations" *Phys. Chem. Chem. Phys.*, **21**, 21761-21775. 2019.

This chapter discusses routes to access electronic excited states within a DFT framework. Specifically, it focuses on the widely used linear-response time-dependent DFT (TDDFT) and orbital optimized DFT (OO-DFT) approaches. The TDDFT equations are presented, along with a discussion about the strengths and weaknesses of the method. The failure of TDDFT for charge-transfer states is used to motivate OO-DFT. The justification behind OO-DFT methods are discussed, as are areas where further development efforts are required. The importance of core-level spectroscopy is also highlighted, along with the challenges faced by standard linear-response methods in modeling core-hole states.

Chapter 10: TDDFT Beyond the Coulson Fischer-Point

The content and figures of this chapter are reprinted or adapted from **Hait, D**.*; Rettig, A.*; Head-Gordon, M. "Beyond the Coulson-Fischer point: characterizing single excitation CI and TDDFT for excited states in single bond dissociations" *Phys. Chem. Chem. Phys.*, **21**, 21761-21775. 2019 ; with permission from the PCCP Owner Societies and all co-authors.

(* indicates equal contribution from authors.)

TDDFT is the most widely used approach for excited state quantum chemistry and is often employed to study photochemical processes. This chapter examines the behavior of the resulting excited state potential energy surfaces beyond the Coulson-Fischer (CF) point in single bond dissociations, when the optimal reference determinant is spin-polarized. Many excited states exhibit sharp kinks at the CF point, and connect to different dissociation limits via a zone of unphysical concave curvature. In particular, the unrestricted $\mathrm{M}_S = 0$ lowest triplet T_1 state changes character, and does not dissociate into ground state fragments. The unrestricted $M_S = \pm 1$ T_1 configuration interaction singles (CIS) states better approximate the physical dissociation limit, but their degeneracy is broken beyond the CF point for most single bond dissociations. On the other hand, the $M_S = \pm 1$ T_1 time-dependent HF (TDHF) states reach the asymptote too soon, by merging with the ground state from the CF point onwards. Use of local exchange-correlation functionals causes $M_S = \pm 1$ T_1 TDDFT states to resemble their unphysical $M_S = 0$ counterpart. The 2 orbital, 2-electron model system of minimal basis H_2 is analytically treated to understand the origin of these issues, revealing that the lack of double excitations is at the root of these remarkable observations. The behavior of excited state surfaces is also numerically examined for species like H_2, NH_3, C_2H_6 and LiH in extended basis sets.

Chapter 11: Square Gradient Minimization

The content and figures of this chapter are reprinted or adapted with permission from **Hait, D.**; Head-Gordon, M. "Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory." *J. Chem. Theory Comput.*, **16**, 1699-1710. 2020. Copyright 2020 American Chemical Society.

This chapter describes a general approach to converge excited state solutions to *any* quantum chemistry orbital optimization process, without the risk of variational collapse. The resulting Square Gradient Minimization (SGM) approach only requires analytic energy/Lagrangian orbital gradients and merely costs 3 times as much as ground state orbital optimization (per iteration), when implemented via a finite difference approach. SGM is applied to both single determinant ΔSCF and spin-purified Restricted Open-Shell Kohn-Sham (ROKS) approaches to study the accuracy of OO-DFT excited states. It is found that SGM can converge challenging states where the Maximum Overlap Method (MOM) or analogues either collapse to the ground state or fail to converge. This chapter also reports that ΔSCF/ROKS predict highly accurate excitation energies for doubly excited states (which are inaccessible via TDDFT). Singly excited states obtained via ROKS are also found to be quite accurate, especially for Rydberg states that frustrate (semi)local TDDFT. The results suggest that excited state OO-DFT methods can be used to push past the limitations of TDDFT to doubly excited, charge-transfer or Rydberg states, making them a useful tool for the practical quantum chemist's toolbox for studying excited states of large systems.

Chapter 12: Restricted Open-Shell Kohn-Sham for Core-Level Excitations

The content and figures of this chapter are reprinted or adapted with permission from **Hait, D.**; Head-Gordon, M. "Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining sub Sub-electronvolt Error from a Restricted Open-Shell Kohn-Sham Approach." *J. Phys. Chem. Lett.*, **11**, 775-786. 2020. Copyright 2020 American Chemical Society.

This chapter presents the use of the SGM algorithm to obtain spin-pure ROKS energies for core excited states of molecules. ROKS/SGM with the modern SCAN/ωB97X-V functionals is found to predict the K edge of C,N,O and F to a root mean squared error of $\sim$0.3 eV. ROKS/SGM is equally effective at predicting L edge spectra of third period elements, provided a perturbative spin-orbit correction is employed. This high accuracy can be contrasted with traditional TDDFT, which typically has greater than 10 eV error and requires translation of computed spectra to align with experiment. ROKS is computationally affordable (having the same scaling as ground state DFT, and a slightly larger prefactor) and can be applied to geometry optimizations/*ab-initio* molecular dynamics of core excited states, as well as condensed phase simulations. ROKS can also model doubly excited/ionized states with one broken electron pair, which are beyond the ability of TDDFT methods.

Chapter 13: Core-Level Spectra of Open-Shell Systems

The content and figures of this chapter are reprinted or adapted from **Hait, D.**; Haugen, E.A.; Yang, Z.; Oosterbaan, K.J.; Leone, S.R.; Head-Gordon, M. "Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations." *J. Chem. Phys.*, **153**, 134108. 2020; with full permission from AIP Publishing and all co-authors.

This chapter utilizes the SGM algorithm to study the effectiveness of OO-DFT at predicting second period element 1s core-level spectra of open-shell systems. Several density functionals (including SCAN, B3LYP and ωB97X-D3) are found to predict excitation energies from the core to singly occupied levels to high accuracy ($\leq$ 0.3 eV RMS error), against available experimental data. Higher excited states are however more challenging by virtue of being intrinsically multiconfigurational. A CI inspired route to self-consistently recouple single determinant mixed configurations from DFT is presented, in order to obtain approximate doublet states with three unpaired electrons. This recoupling scheme is used to predict the C K-edge spectra of the allyl radical, the O K-edge spectra of CO^+ and the N K-edge of NO_2 to high accuracy relative to experiment, indicating substantial promise in using this approach for computation of core-level spectra for doublet species (vs TDDFT, EOM-CCSD or using unrecoupled mixed configurations). General guidelines for computing core-excited states from OO-DFT are also presented.

Chapter 14: Afterword on Electronic Excited States

This chapter summarizes the results of the preceding four chapters. SGM is placed in the context of other OO algorithms developed before and after. The strengths of OO-DFT approaches is described, along with an application for studying the transient C K-edge absorption spectrum of strong-field ionized ${CCl_4}^+$. Finally, connections between OO-DFT and related wavefunction approaches is discussed, as well as the role OO-DFT might play with regards to DFA development. It is hoped that OO-DFT would see wider use in future to model states that are challenging for (or inaccessible with) TDDFT.

2.1 Works Not Discussed in This Dissertation

Several scientific works that the author has contributed are not described in detail in this dissertation. A list of such works is however provided, to give a more complete account of the odyssey. Several others remain on the pipeline, and will be published in the near future. * indicates equal contribution from authors.

Excited States

1. Cunha, L.A.*; **Hait, D**.*; Kang, R.; Mao, Y. Head-Gordon, M. "Relativistic Orbital Optimized Density Functional Theory for Accurate Core-Level Spectroscopy. *J. Phys. Chem. Lett.,* **13**, 3438–3449. 2022.

2. Oosterbaan, K.J.; White, A.F.; **Hait, D**.; Head-Gordon, M. "Generalized Single Excitation Configuration Interaction: An Investigation into the Impact of the Inclusion of Non-Orthogonality on the Calculation of Core-Excited States."*Phys. Chem. Chem. Phys.*, **22**, 8182-8192.

Errors in DFT

1. **Hait, D**.; Head-Gordon, M. "xDH double hybrid functionals can be qualitatively incorrect for non-equilibrium geometries: Dipole moment inversion and barriers to radical-radical association using XYG3 and XYGJ-OS." *J. Chem. Phys.*, **148**, 171102. 2018.

2. **Hait, D**.; Head-Gordon, M. "Delocalization errors in density functional theory are essentially quadratic in fractional electron number." *J. Phys. Chem. Lett.*, **9**, 6280-6288. 2018.

Selected Configuration Interaction

1. **Hait, D.**; Tubman, N.M.; Levine, D.S.; Whaley, K.B.; Head-Gordon, M. "What levels of coupled cluster theory are appropriate for transition metal systems? A study using near exact quantum chemical values for 3d transition metal binary compounds." *J. Chem. Theory Comput.*, **15**, 5370-5385. 2019.

2. Levine, D.S.; **Hait, D**.; Tubman, N.M.; Lehtola, S.; Whaley, K.B.; Head-Gordon, M. "CASSCF with Extremely Large Active Spaces using the Adaptive Sampling CI Method." *J. Chem. Theory Comput.*, **16**, 2340-2354. 2020.

3. Tubman, N.M. *et.al.* "Modern Approaches to Exact Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method." *J. Chem. Theory Comput.* **16**, 2139-2159. 2020.

4. Eriksen, J.J. *et.al.* "The Ground State Electronic Energy of Benzene." *J. Phys. Chem. Lett.*, **11**, 8922-8929. 2020.

Single reference wavefunction theory

1. Rettig, A.*; **Hait, D**.*; Bertels, L.W.; Head-Gordon, M. "Third order Møller-Plesset theory made more useful? The role of density functional theory orbitals" *J. Chem. Theory Comput.,* **16**, 7473-7489. 2020.

Catalysis Applications

1. Fang, J.; **Hait, D**.; Head-Gordon, M.; Chang, M.C.Y."Chemoenzymatic platform for synthesis of chiral organofluorines based on type II aldolases." *Angew. Chem. Int. Ed.*, **58**, 11841-11845. 2019.

2. Witzke, R.J.; **Hait, D.**; Chakarawet, K.; Head-Gordon, M.; Tilley, T.D. "Bimetallic mechanism for alkyne cyclotrimerization with a two-coordinate Fe precatalyst." *ACS Catal.,* **10**, 7800–7807. 2020.

3. Witzke, R.J.; **Hait, D.**; Head-Gordon, M.; Tilley, T.D. "Two-Coordinate Iron(I) Complexes on the Edge of Stability: Influence of Dispersion and Steric Effects." *Organometallics,* **40,** 1758–1764. 2021.

Part I

Electrical Response Properties

Chapter 3

Foreword to Electrical Response Properties

But in the quietness of thought:
Me this unchartered freedom tires;
I feel the weight of chance-desires:
William Wordsworth. "Ode to duty"

3.1 The Density Problem

The last four decades have seen significant and consistent improvements in DFA quality, in terms of predicting ground state energy differences [55, 115]. While there were some attempts to also consider the accuracy of electron density in DFA development [118, 119], such efforts were mostly abandoned by 2005 as energies emerged as the exclusive route for assessing and empirically training DFAs. This development was in many cases made possible by development of highly accurate benchmark datasets from wavefunction theory [56], which led to thousands of distinct chemically relevant datapoints that could be utilized for functional development [55, 111]. In particular, this large volume of data was used to train a number of DFAs, which exploited the vastness of the mGGA space to create rather complex functional forms (with $25-60$ fitted parameters) that resulted in low training set error [101, 120–127]. Other efforts focused on selecting terms that led to maximum transferability between training and test sets, resulting in forms with < 15 fitted parameters [63, 96, 108]. In virtually all cases, it was implicitly assumed that the densities would improve simultaneously with the energies, as developed DFAs approached the exact functional.

In 2017, Ref 115 challenged such ideas by comparison between densities from DFAs and CCSD for atomic systems (like Ne or N^{3+}). The results therein suggested that for functionals developed between 1980 and ~ 2000, the more recent DFAs tended to yield better densities, reflecting steady advancement towards the exact ideal. However, from the mid 2000s onward, many functionals with very large density errors (worse than LSDA in fact) were developed, resulting in the trendline for density errors vs year of DFA development to go up instead of down. Indeed, it appeared that errors in density had been effectively decoupled from errors in energy. The authors highlighted these results with the provocative title of "Density functional theory is straying from the path toward the exact functional". This work received considerable attention from the scientific community [116, 128–130], setting off a raging debate about the need for new strategies for DFA development.

There were aspects of Ref 115 that received criticism [130]. There were doubts expressed as to whether highly charged cations (or even atomic species) studied in that work were representative of chemical systems of broad interest [116], whether CCSD was sufficiently accurate for comparisons [130] and whether the conclusions were fair to all modern functionals [130]. Indeed, the worst performers reported in Ref 115 were highly parameterized functionals (30+ parameters) developed with the same general strategy of using all available data for training and none for testing transferability [121, 123–125]. The removal of those potentially overfitted DFAs from Ref 115 would indicate that density quality merely ceased to significantly improve with modern functionals, instead of worsening. Regardless of these aspects, it cannot be denied that Ref 115 forced a major reconsideration about approaches functional development and catalyzed a large body of new work [50, 131–135]. It quickly became clear that it was necessary to evaluate both the accuracy of existing DFAs for molecular systems of chemical interest, as well as potentially develop datasets of density related quantities that can be used in conjunction with energies for functional development [116].

Many of the arguments made in favor of utilizing densities in functional development are philosophical, rooted in the argument that DFT should yield the right answer for the right reasons (exact functional on exact density to yield exact energy), and not accidentally yield the right answer for the wrong reasons (inexact functional mapping an unacceptable density to an acceptable energy) [116, 130]. This is of course completely correct, but at times is brushed aside by 'pragmatists' on a quest for improving energy prediction quality. However, as we will show in the next section, it is impossible to disentangle accuracy in predicting electron densities from accuracy in predicting energies for chemical applications beyond isolated gas phase molecules. The electron density modulates the response of the electronic energy to external potentials from the environment (be it an electromagnetic field or other molecules) and the pragmatic course therefore is to simultaneously aim for both improved densities and improved energies, through the use of density dependent datasets in functional development. This should help lower both the possibility of overfitting to only one form of training data, as well as prevent minimally parameterized models from being underconstrained to the point of being unable to predict densities.

3.2 Connecting Densities to Energy

Let us consider some chemical system with the electronic Hamiltonian $\mathbf{H}_e$, ground state density $\rho(\vec{r})$ and ground state energy $E\left[\rho(\vec{r})\right]$. Let us subsequently define a perturbed family of Hamiltonians given by $\mathbf{H}'(\lambda, \vec{r'})$ defined to be:

$$\mathbf{H}'(\lambda, \vec{r'}) = \mathbf{H}_e + \lambda\delta(\vec{r} - \vec{r'}) \tag{3.1}$$

$\delta(\vec{r} - \vec{r'})$ is a local one-particle potential, and therefore the HK theorems apply to $\mathbf{H}'(\lambda, \vec{r'})$. Let the ground state density be $\rho_{\lambda,\vec{r'}}(\vec{r})$. The ground state energy of $\mathbf{H}'(\lambda, \vec{r'})$ is then $E[\rho_{\lambda,\vec{r'}}(\vec{r})]$. From first order perturbation theory in λ, we know that:

$$\left(\frac{\partial E[\rho_{\lambda,\vec{r'}}(\vec{r})]}{\partial\lambda}\right)_{\lambda=0} = \rho(\vec{r'}) \tag{3.2}$$

yielding the density for an arbitrary point $\vec{r'}$.

This protocol might appear to be overly convoluted, seeing that evaluation of E requires knowledge of ρ. However, it is important for two reasons. The first applies to cases where E is not an explicit functional of ρ, but a rather more complicated implicit functional. Specifically, the Slater determinant density ρ_{KS} for rung 5 functionals can (and generally is) distinct from the ground state density ρ, due to the presence of wavefunction correlation. The above protocol can therefore be used to find the density for such problems, and in fact could be generalized to compute densities for any arbitrary quantum chemistry method.

The second reason is that the connection above clearly shows that there is no functional that can yield the exact energy with an incorrect density for arbitrary external potentials. In other words, a functional that might accidentally yield the exact energy for the $\lambda = 0$ case with an incorrect density would immediately fail for nonzero λ. The density therefore modulates the response of the energy to external local potentials (which are trivially decomposable into a sum of Dirac δ functions). Since electrons are often subjected to local potentials beyond electron-nuclear attraction (such as in spectroscopy, electrochemistry or hybrid calculations with classical charge distributions present), a correct description of all interactions necessitates functionals that yield both good energies and good densities.

3.3 Multipole Moments

Assessing the accuracy of electron densities is a somewhat trickier task than energies, as the former is a function and the latter a scalar. It is possible to compare densities over a large grid of points, or utilize statistical distance measures like the Kullback–Leibler divergence [136]. However, evaluating the exact density (or a sufficiently good correlated wavefunction theory approximation thereof) for a large grid of points is computationally cumbersome.

An alternative is to utilize molecular multipole moments, which are spatial moments of the total charge density of a molecule. For a molecule with electron number density $\rho(\vec{r})$ and nuclei of atomic number $\{Z_A\}$ at positions $\{\vec{R}_A\}$, the total charge density is:

$$P(\vec{r}) = -\rho(\vec{r}) + \sum_A Z_A \delta(\vec{r} - \vec{R}_A) \tag{3.3}$$

The general multipole moment component M_{abc} is then defined to be:

$$M_{abc} = \int x^a y^b z^c P(\vec{r}) d\vec{r} \tag{3.4}$$

The order of the multipole moment is given by $a+b+c$, and all moments of the same order collectively constitute a tensor of that rank. The lowest order moment tensors are:

1. Charge (zeroth order):

$$M_{000} = -\int \rho(\vec{r}) d\vec{r} + \sum_A Z_A \tag{3.5}$$

2. Dipole moment vector $\vec{\mu}$ (first order), defined by

$$\vec{\mu} = M_{100}\hat{x} + M_{010}\hat{y} + M_{001}\hat{z} = -\int \vec{r}\rho(\vec{r}) d\vec{r} + \sum_A Z_A \vec{R}_A \tag{3.6}$$

3. Second order: Quadrupole moment tensor (second order):

$$\begin{aligned}\mathbf{Q} &= M_{200}\left(\hat{x}\otimes\hat{x}\right) + M_{020}\left(\hat{y}\otimes\hat{y}\right) + M_{002}\left(\hat{z}\otimes\hat{z}\right) + M_{110}\left(\hat{x}\otimes\hat{y} + \hat{y}\otimes\hat{x}\right) \\ &\quad + M_{101}\left(\hat{x}\otimes\hat{z} + \hat{z}\otimes\hat{x}\right) + M_{011}\left(\hat{y}\otimes\hat{z} + \hat{z}\otimes\hat{y}\right) && (3.7) \\ &= -\int \left(\vec{r}\otimes\vec{r}\right)\rho(\vec{r}) d\vec{r} + \sum_A Z_A \left(\vec{R}_A \otimes \vec{R}_A\right) && (3.8)\end{aligned}$$

where the dyadic product of the position vector $\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is $\vec{r}\otimes\vec{r} = \begin{pmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{pmatrix}$

In general, a multipole moment is composed of an integral of the electron density times some spatial component, plus a nuclear component that is trivially computable under the fixed nucleus assumption of the Born-Oppenheimer approximation. They therefore can act as simple, global measures of the accuracy of $\rho(\vec{r})$. They are also physically meaningful, as they describe the response of the electronic energy to spatially polynomial potentials. In other words, if we perturb the electronic Hamiltonian $\mathbf{H}_e$ with a potential $\lambda x^a y^b z^c$, we find that the resulting ground state energy is $E(\lambda) = E + \lambda M_{abc} + O(\lambda^2)$, where E is the unperturbed ground state energy. Multipole moments therefore appear to be easy to evaluate quantities that will yield information both about the accuracy of the electron density as well as the physical response of the electronic energy to external potentials.

One caveat regarding their use is that multipole moments in general are not translationally invariant (i.e. their value depends on where the origin of the system of coordinates is defined). Only the lowest non-zero order moment is independent of the choice of the origin. For charged systems therefore, neither the dipole nor any higher order multipole moments are translationally invariant. However $\vec{\mu}$ is origin independent for neutral molecules, leading to widespread use in many areas of chemistry. Higher order moments can however be converted to translationally invariant cumulants of the electron density (as described in Chapter 5), which permits their use for assessing densities.

3.4 Multipole Polarizabilities

Multipole moments only describe the lowest order response of the energy to perturbing electric fields. However, sometimes the potentials are strong enough for higher order responses to become relevant. For instance, if a molecule is subjected to a spatially constant electric field $\vec{\mathcal{E}} = \mathcal{E}_x\hat{x} + \mathcal{E}_z\hat{y} + \mathcal{E}_z\hat{z}$, it experiences a potential $V(\vec{r}) = x\mathcal{E}_x\hat{x} + y\mathcal{E}_z\hat{y} + z\mathcal{E}_z\hat{z}$. The energy of the molecule can then be described as a Taylor series of the form:

$$E(\vec{\mathcal{E}}) = E(0) - \vec{\mu}\cdot\vec{\mathcal{E}} - \frac{1}{2!}\mathbf{Tr}\left[\boldsymbol{\alpha}\left(\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\right)\right] - \frac{1}{3!}\mathbf{Tr}\left[\boldsymbol{\beta}\left(\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\right)\right] - \frac{1}{6!}\mathbf{Tr}\left[\boldsymbol{\gamma}\left(\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\otimes\vec{\mathcal{E}}\right)\right]\ldots \tag{3.9}$$

where the second rank tensor $\boldsymbol{\alpha}$ is called the polarizability, the third rank tensor $\boldsymbol{\beta}$ is called the first hyperpolarizability, the fourth rank tensor $\boldsymbol{\gamma}$ is called the second hyperpolarizability, and so on. In general, (hyper)polarizabilities are second and higher order derivatives of the energy with respect to the strength of the electric field. Furthermore, electric fields lead to mixing between the (field-free) ground and excited electronic states of the molecule. This is not evident at first order perturbation theory for the energy, and therefore the multipole moments only convey information about the unperturbed electron density. The higher order derivatives however control how the field alters the density of the molecule and thereby depend on information about the excited state spectrum. They are therefore of relevance to nonlinear optics. We note that Eqn 3.9 has no time dependence, and the associated (hyper)polarizabilities are called static (hyper)polarizabilities. The frequency dependent variants of those tensors are useful for describing the response to time-dependent fields.

More directly, multipole polarizabilities are important for correctly describing the energy and densities of molecules interacting with strong fields. The exact functional should be able to describe such behavior perfectly, as any external electric field only results in a local one body potential of the type supported by the HK theorems. Information about multipole polarizabilities thus adds to the body of data that is directly useful for DFA development.

3.5 Objectives

In this half of the dissertation, we will study the ability of DFAs to predict electrical response properties like molecular multipole moments and polarizabilities. Specifically, we investigate the dipole moment in Chapter 4, the second cumulant of the electron density (derived from the quadrupole moment) in Chapter 5 and the static dipole polarizability in Chapter 6. We note that the first and third have been studied earlier (as noted in the respective chapters) but not comprehensively and certainly not with modern DFAs. We therefore present benchmark datasets for all three properties at the CCSD(T)/CBS limit (where CBS stands for the complete basis set limit, indicating that the results are converged with respect to spatial degrees of freedom available to the electrons) for large numbers (≥ 100) of small, neutral, molecular systems. The resulting datasets are of sufficient accuracy (vs FCI) for training and assessment of DFAs. We utilize them to study the performance of popular and recent DFAs to gauge whether or not modern DFA development has strayed far from 'the true path'. Specifically, we examine whether highly parameterized mGGAs (or their hybrid equivalents) have been overparameterized on account of the 'unchartered freedom' offered by the fitting process. We also study whether some minimally parameterized models are underconstrained to deal with data outside their original training and fitting sets of molecular ground state energy differences. Our results reveal that while performance improves along ascending Jacob's ladder (in that the best functional in each rung surpasses the best in the rungs below), modern DFAs (especially mGGAs and derived hybrids) aren't particularly good at predicting these electrical response properties. Use of these datasets for future functional development would therefore be useful towards yielding DFAs that can model a broader range of applied local potentials beyond just bare V_{ne}.

During the course of our investigations, we also noted that these properties were rather sensitive probes for electronic structure, often highlighting catastrophes that would not be immediately apparent for energies. For instance, the dipole moment along the dissociation curves of charged [76, 95], and polar [78–81] molecules often reveal massive delocalization driven challenges. However, nonpolar bond dissociation in neutral species was generally thought to be well described by DFAs as there are no driving forces for fractional charges. We discovered that a number of DFAs preserve a 'partial bond' for H_2 even at long stretches, which is clearly unphysical. This at times manifests via a small, spurious barrier to dissociation but is hard to immediately gauge from the potential energy surface. The static dipole polarizability however lays the problem bare, offering a very clear picture of where the DFA predicted electronic structure is problematic. Chapter 7 discusses this use of static polarizability as a computational probe in greater detail, showing where DFAs fail to describe homoyltic dissociations of nonpolar bonds.

This half of the dissertation therefore covers both quantitative deviations from exact behavior at equilibrium geometries (where behavior of the electrical response properties is qualitatively acceptable) and catastrophic failure at non-equilibrium nuclear configurations.

Chapter 4

Dipole Moments

4.1 Introduction

Density functional theory (DFT) is currently the most popular approach for calculating the electronic structure of molecules and extended materials [55, 90, 137]. Although DFT is formally exact [41], the true functional that maps electron density to electronic energy remains unknown. Most practical calculations thus employ various density functional approximations (DFAs) within the Kohn–Sham [48] DFT framework. Hundreds of DFAs have consequently been proposed over the last few decades, with development strategies focusing on empirical fitting to large benchmark datasets [101, 120, 126], satisfaction of various exact physical constraints [54, 57, 138] or intermediate routes combining aspects of both approaches [62, 96]. With rare exceptions [118, 119, 139, 140], fitting based approaches have historically favored recovering accurate energies over accurate densities, in the hope that DFAs predicting improved energies were better approximations to the exact functional, and would thus automatically yield more accurate electron densities.

This approach has recently been questioned by studies that compared coupled cluster singles and doubles (CCSD) electron densities of a few atomic [115] and diatomic [131] species to those predicted by various DFAs. These studies appear to suggest that many modern functionals (that are excellent for energetics) are unable to predict very good densities, deviating from an earlier trend of DFAs simultaneously improving energies and densities. This has led to suggestions that DFT has strayed from the path to the exact functional [115], causing lively discussions in the scientific community [50, 116, 128, 130, 132–135, 141–143]. Discussions about "true paths" aside, accurate electronic densities are necessary not only for intellectual completeness, but also to quantitatively predict the response of the electronic energy to external potentials (such as ones in spectroscopic experiments, or intermolecular interactions or at an electrode). It is therefore desirable to develop functionals that would simultaneously predict the energy and density to high levels of accuracy, perhaps by incorporating density information as part of functional development.

Multipole moments describe spatial moments of the electron density, and the accuracy with which they can be calculated is an integrated measure of the quality of the corresponding density. The simplest possible multipole is the dipole $\vec{\mu} = \int \vec{r}P(\vec{r})d\vec{r}$ (where $P(\vec{r})$ is *charge* density), which is the sum of the first spatial moment of the electron density and a nuclear charge contribution. $\vec{\mu}$ is not a perfect stand-in for the density (not least because it can vanish due to symmetry) and it is certainly possible for an incorrect density to accidentally yield a correct value of $\vec{\mu}$. However, $\vec{\mu}$ is perhaps the simplest observable that captures errors in the underlying density. $\vec{\mu}$ also determines the response of the energy to a constant electric field and is relevant to long-range electrostatics in classical force-fields [144–146], making it a relevant density derived quantity to examine for DFA testing and development.

Although there have been previous studies examining the accuracy of DFT in estimating vibrationally averaged experimental dipole moments [141, 147], there does not appear to be any large dataset of benchmark equilibrium $\vec{\mu}$. The distinction between experimental and equilibrium $\vec{\mu}$ is important as they are known to differ by 3% or more in many cases like NH_3 [148]. The accuracy of electronic structure approximations (in the absence of vibrational corrections) in directly reproducing experimental dipole moments thus may not be indicative of the quality of the approximation. This led us to develop a dataset of 200 coupled cluster dipole moments for the purpose of assessing and developing density functionals. The main constituent of our dataset is an estimate of the unrestricted CCSD(T) [23] dipole moment at the complete basis set (CBS) limit for 95 normal valent systems (88 molecules and 7 non-covalent complexes, containing up to 3 non-hydrogen atoms and with all atoms at normal valence) and 57 incomplete valence and open-shell systems (50 molecules and 7 non-covalent complexes where not all atoms were at normal valence) at their equilibrium geometries (given in the Supporting Information). For simplicity, we only report magnitudes $\mu = |\vec{\mu}|$ for most species, but we do supply the direction for certain cases like CO, where different DFAs predict different polarities. We also assessed the performance of 88 popular or recent density functionals [54, 57–59, 61–63, 89, 91, 92, 94, 96, 97, 101, 107, 108, 118–127, 138, 149–190] in predicting equilibrium geometry μ over our dataset, in order to gauge the relative accuracy of these functionals over a substantially larger dataset than those employed by previous studies [115, 131]. The remaining data-points are $\vec{\mu}$ along the potential energy surfaces (PES) of FH and FCl, enabling exploration of some non-equilibrium configurations.

4.2 Computational Methods

All the calculations were done using a development version of Q-Chem 4 [191]. The equilibrium geometries employed to estimate $\vec{\mu}$ for each of the 152 species considered were obtained from a variety of sources and no geometry re-optimization was done for individual methods. Experimental equilibrium geometries (obtained from the NIST Computational Chemistry Database [192] with two exceptions: BH_2F [193] and BH_2Cl [194]) were employed whenever possible for the molecular species. Molecular geometries were optimized at the MP2/cc-pVTZ level of theory when experimental geometries were unavailable, as MP2 equilibrium geometries in general are known to be sufficiently accurate [195]. The geometries of the non-covalently bounded complexes were taken from the NCCE31 [196] (for closed-shell species) and TA13 [197] (for open-shell species) databases. All the molecular geometries employed in the present study are provided in the Supporting Information .

Hartree-Fock (HF) $\vec{\mu}$ were obtained from spin unrestricted calculations with the aug-cc-pCV5Z [70, 198–200] basis, which appeared to have essentially converged to the CBS limit as they only showed a 0.04% RMS deviation versus the equivalent aug-cc-pCVQZ numbers. Stability analysis was performed at the aug-cc-pCVQZ level to ensure all solutions were at a minima. Spin unrestricted DFT calculations were done with the aug-pc-4 [65–69] basis for functionals spanning Rungs 1-4 in Jacob's Ladder and it was assumed that such 5ζ basis results ought to be essentially at CBS as well (though this may not strictly be true for a few ill-behaved functionals, as discussed in the results section). Local xc integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points for all atoms, while nonlocal VV10 correlation was calculated over an SG-1 [201] grid (which consists of a subset of points employed in a grid with 50 radial and 194 angular points). Stability analysis was done at the aug-pc-2 level to determine which SCF solutions were potentially unstable, and the problematic aug-pc-4 cases were reoptimized to ensure that they were at a minima.

Dipole moments for correlated wavefunction methods (MP2/CCSD/CCSD(T)) were obtained from two point central finite differences (using an electric field strength of 10^{-4} a.u.). The CBS limit was obtained by extrapolating the correlation component of μ from aug-cc-pCV5Z and aug-cc-pCVQZ results for the smaller species; and from the aug-cc-pCVQZ and aug-cc-pCVTZ results for the remainder of the dataset. The extrapolation was done via the two point formula $\mu_n^{\text{corr}} = \mu_\infty^{\text{corr}} + A/n^3$ used in Ref 202, which was found to be accurate to $\approx 0.2\%$ in predicting aug-cc-pCV5Z dipoles of some of the smaller species from aug-cc-pCVQZ and aug-cc-pCVTZ results, indicating that it was reasonably accurate. All the calculations were done using unrestricted orbitals (unless specified otherwise) and no orbitals were held frozen.

All calculations with double hybrid functionals for species at equilibrium geometry were done with restricted orbitals on account of N-representability violations in unrestricted MP2 [203]. Dipole moments were calculated using the same finite difference approach as correlated wavefunction methods. The non-MP2 component of $\vec{\mu}$ was assumed to have essentially converged to CBS at the aug-cc-pCV5Z level, while the MP2 correlation component was extrapolated to CBS from aug-cc-pCVQZ and aug-cc-pCVTZ results, using the same formula as above. The xc integrals for double hybrid functionals were calculated using the same grids as all other functionals. Calculations employing the XYGJ-OS [189] functional were accelerated using the RI approximation [204] with the riMP2-cc-pVQZ (riMP2-cc-pVTZ) auxiliary basis [205, 206] for aug-cc-pCVQZ (aug-cc-pCVTZ) calculations.

Wavefunction theory (WFT) calculations for the FH and FCl PES's were performed in the same manner as equilibrium calculations, with the CBS limit for correlated methods found from two point extrapolation from aug-cc-pCVQZ and aug-cc-pCVTZ results. Double hybrid calculations were performed similarly to equilibrium calculations as well, with the major differences being use of unrestricted orbitals (since restricted references are grossly inadequate in describing bond dissociations [14]) and treatment of aug-cc-pCVQZ numbers as the CBS limit for the non-MP2 component instead of aug-cc-pCV5Z results. Other DFT calculations were done with the aug-pc-3 basis, with the same integration grids as the equilibrium cases. Stability analysis was performed on all SCF solutions to ensure the determinant was a minimum. No expensive aug-pc-4/aug-cc-pCV5Z calculations were performed for any of the functionals, as the basis set incompleteness errors at the quadruple ζ level are small compared to the qualitative questions that the exploration along those potential energy curves are addressing.

The error in μ against the reference value μ_{ref} is defined to be $\dfrac{\mu - \mu_{\text{ref}}}{\max(\mu_{\text{ref}}, 1\text{ D})} \times 100\%$. This regularized form was chosen to ensure that the errors were relative for large μ_{ref} (to prevent relatively small perturbations of density in ionic species from dominating) but were absolute for small μ_{ref} (to prevent a too small denominator from skewing the analyis). This effectively results in a regularized relative error as a majority of the database has $\mu_{\text{ref}} \geq 1$ D. The choice of 1 D as the cutoff is somewhat arbitrary, but it allows a seamless transition from absolute error (for small μ_{ref}) to relative error (for large μ_{ref}). Purely absolute and purely relative errors are also supplied in the Supporting Information, with trends in the former being consistent with the trends in our regularized relative error. Purely relative errors are however considerably less useful, as species with < 0.1 D (such as the OF radical) wind up heavily dominating total error and the rest barely make a significant contribution.

4.3 Results for Equilibrium Geometries

We divided the 152 species in our dataset into two groups: 81 where the stable HF solution did not have spin-polarization, and 71 where the stable HF solution was spin-polarized. The former subset is henceforth referred to as NSP (not spin-polarized, or spin-restricted) and the latter as SP (spin-polarized, or unrestricted). Conventional electronic structure methods like CCSD, MP2 or the various DFAs are unlikely to experience catastrophic failure over the NSP subset, but the SP species have the potential to be more challenging.

The regularized relative error metric defined in the preceding section was used to evaluate the performance of HF, MP2, CCSD and 76 DFAs against the CCSD(T) benchmark. The errors for a selection of popular and new DFAs are given in Table 4.1, while Fig 4.1 provides a visual representation of performance for a subset of the methods. The full list of all errors for all methods is provided in the Supporting Information, along with all computed dipoles.

4.3.1 Performance of Wavefunction Theory

As expected, CCSD is the most effective wavefunction theory (WFT) in reproducing the CCSD(T) benchmark, giving a regularized RMSE of 3.95% over the whole dataset. This difference is large enough, however, to clearly show that CCSD itself is not a suitable replacement, and, by implication, that CCSD densities themselves have limitations. Furthermore,

Method	Class	RMSE			MAX	ME
		All	NSP	SP		
CCSD	WFT	3.95	3.01	4.8	12.87	1.86
MP2	WFT	37.45	3.96	54.63	367.72	7.97
RMP2	WFT	7.32	3.96	9.84	32.49	0.55
HF	WFT	18.61	12.81	23.53	57.35	7.18
RHF	WFT	20.1	12.81	26.03	76.64	9.67
SVWN1RPA	LSDA	13.67	8.93	17.59	66.22	-2.13
Slater	LSDA	14.82	12.11	17.4	49.31	-3.95
N12	GGA	8.85	5.13	11.74	45.29	-1.72
B97-D	GGA	9.87	7.95	11.68	43.3	-3.48
PBE	GGA	11.89	9.55	14.09	45.02	-4.15
BLYP	GGA	12.08	8.55	15.13	55.82	-4.16
SOGGA11	GGA	21.44	11.31	28.95	197.74	-2.35
mBEEF	mGGA	7.56	6.55	8.56	29.97	-2.45
SCAN	mGGA	8.44	6.8	9.98	31.81	-1.31
MS2	mGGA	8.94	7.76	10.12	28.98	-2.83
TPSS	mGGA	9.93	7.92	11.82	40.51	-3.02
revM06-L	mGGA	10.68	10.86	10.46	34.1	-1.81
M11-L	mGGA	11.05	11.01	11.09	34.73	-5.65
B97M-V	mGGA	11.61	8.44	14.39	71.08	-3.81
M06-L	mGGA	12.09	9.17	14.73	79.4	-2.42
MN15-L	mGGA	12.19	10.37	13.99	54.08	-2.6
MN12-L	mGGA	53.24	7.59	77.47	645.34	2.93
SOGGA11-X	hGGA	4.84	3.78	5.82	23.93	0.69
ωB97X-V	hGGA	5.07	4.32	5.8	18.82	1.81
PBE0	hGGA	5.18	4.52	5.85	22.25	-0.32
B97-2	hGGA	5.34	4.5	6.16	26.92	-0.58
HSEHJS	hGGA	5.35	4.56	6.14	22.64	-0.45
ωB97X-D	hGGA	5.36	4.43	6.25	20.8	1.35
camB3LYP	hGGA	5.83	3.68	7.56	40.23	1.7
B97	hGGA	6.06	4.99	7.09	29.03	-1.23
CAM-QTP01	hGGA	6.88	4.94	8.57	36.16	3.62
B3LYP	hGGA	6.98	4.43	9.05	46	-0.82
rCAM-B3LYP	hGGA	7.48	5.43	9.27	35.23	4.18
HFLYP	hGGA	19.53	14.87	23.76	66.79	9.11
PW6B95	hmGGA	4.87	3.76	5.88	24.18	-0.44
M05	hmGGA	4.95	4.51	5.4	18.47	-0.14
ωM05-D	hmGGA	4.96	4.46	5.49	17.96	0.87
M06	hmGGA	5.69	5.69	5.68	19.27	-0.63
ωB97M-V	hmGGA	5.74	4.1	7.16	31.94	0.18
SCAN0	hmGGA	6.28	4.46	7.86	22.56	1.17
MS2h	hmGGA	7.08	6.13	8.02	24.82	-1.58
TPSSh	hmGGA	7.42	5.96	8.79	31.94	-1.63
M06-2X	hmGGA	7.94	3.76	10.9	67.07	1
MN15	hmGGA	10.53	4.05	14.78	79.28	0.45
BMK	hmGGA	11.1	3.59	15.77	125.58	1.51
M11	hmGGA	13.06	5.51	18.18	135.49	4.67
M06-HF	hmGGA	18.13	13.15	22.5	110.82	6.17
MN12-SX	hmGGA	27.11	5.26	39.26	300.82	1.92
DSD-PBEPBE-D3	dhGGA	3.76	2.64	4.71	20.41	0.1
XYGJ-OS	dhGGA	4.19	2.17	5.68	19.41	-0.11
B2GPPLYP	dhGGA	4.36	2.38	5.84	31.58	0.54
ωB97X-2	dhGGA	4.52	3.31	5.6	22.37	0.87
B2PLYP	dhGGA	5.31	2.98	7.08	36.19	-0.39
PWPB95-D3	dhmGGA	3.61	2.52	4.54	18.14	-0.29
PTPSS-D3	dhmGGA	4.74	3.62	5.76	19.33	-0.77

Table 4.1: RMS regularized errors (RMSE), Mean regularized errors (ME) and Maximum absolute regularized errors (MAX) for selected electronic structure methods over the 152 species at equilibrium, expressed as percentages. The RMSE of the spin-polarized (SP) and non-spin polarized (NSP) subsets of the dataset are also reported separately. The functional with the lowest and highest cumulative RMSE in each class is reported, along with other widely used functionals within the same class. WFT stands for wavefunction theory, LSDA for local spin density approximation, GGA for generalized gradient approximation, mGGA for meta-GGA, hGGA for hybrid GGA, hmGGA for hybrid meta-GGA, dhGGA for double hybrid GGA and dhmGGA for double hybrid meta-GGA.

there exist cases like the PS radical where CCSD deviates from CCSD(T) by as much as 13%, raising some questions about whether CCSD(T) itself is adequate as a benchmark for those species (as it only adds a perturbative triples correction to CCSD). However, comparison between CCSD(T) and the more robust (and computationally expensive) CCSD(2) [207] method at the aug-cc-pCVTZ level for species where the CCSD/CCSD(T) regularized deviation is $> 6\%$ reveals that CCSD(2) and CCSD(T) are generally within 1% of each other (full table given in Supporting Information), indicating that the latter is adequate for our purposes. Furthermore, comparison to available experimental values for non-hydrogen containing species (where vibrational averaging is expected to have the least impact) indicates no glaring issues with the CCSD(T) benchmark. These two consistency checks thus collectively suggest that CCSD(T) is likely to be accurate to $\approx 1\%$ even in the worst of cases, and is probably much more accurate for better behaved species, allowing it to serve as a benchmark for our purposes.

The surprisingly poor performance of MP2 is a consequence of N-representability violation in unrestricted MP2 [203], which leads to some very poor numbers for some even-electron singlet systems (like HPO) where spin symmetry is broken, as well as for radicals like ClO_2. The species in the spin-unpolarized NSP subset do not cause much trouble for MP2, and the resulting regularized RMSE of 3.96% is very close to the CCSD RMSE of 3.01% over this subset. On the other hand, the massive 54.63% RMSE over the more challenging spin-polarized SP subset indicates that unrestricted MP2 is a poor choice for estimating μ for such species. Restricted MP2 (RMP2) however performs quite well, giving a much better RMSE of of 9.84% over the SP subset, despite operating on a restricted reference that is higher in energy than the symmetry broken solution. RMP2 consequently has an overall RMS error of 7.3% over the whole dataset, which, as we shall see, is not too bad relative to the best hybrid DFAs.

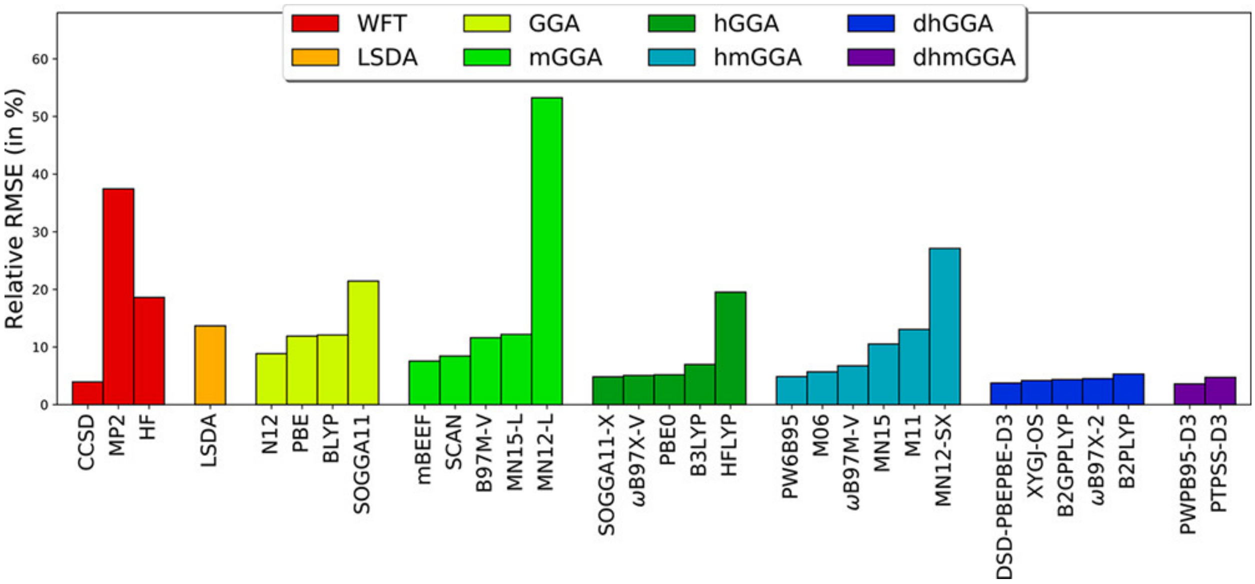


Figure 4.1: Visual representation of the RMSE for a selection of methods listed in Table 4.1.

Not unexpectedly, HF theory performs least well out of all the WFT methods, giving a regularized RMS error of 18.61%. Using purely restricted HF does not however lead to a significant deterioration in quality, as it only adds $\approx 1.5\%$ to the regularized RMS error. It is well known that neglect of correlation in HF theory typically leads to systematic overestimation of dipole moments. Qualitatively, this can be viewed as arising from the fact that antibonding orbitals, which are partly occupied due to correlation, but are empty at the HF level, have polarization opposite to bonding orbitals.

4.3.2 Performance of DFT

There are a number of general features that emerge from inspection of Table 4.1 that are likely to be transferable beyond our specific dataset. First, it is striking that the total RMSE for the best DFA in each class decreases as one ascends Jacob's ladder. This is a nice validation of the fact that there is additional physical content at each level of the ladder. If that additional physical content is used effectively, it can contribute to higher levels of accuracy in the dipole moments, and by inference, the electron density. By contrast, it is also striking that the worst GGA, mGGA and hybrid functionals actually have larger cumulative RMSE than the LSDA. From this we can infer that the additional physical content contained at each higher level of the ladder also leads to greater flexibility in the design of the functional, which, indeed, can degrade the accuracy of the dipole moments relative to LSDA which is a simple parameter-free model. This is consistent with the general conclusions in Refs 115131, which also found that the best performing hybrid functionals yielded superior density predictions than the best local functionals, while several modern functionals (both hybrid and local) gave performance comparable to or worse than LSDA. The overall relative performance of functionals in predicting dipole moments was thus not too dissimilar to the performance in predicting density [115, 131], despite the differences in methodology.

The best functionals for dipole moment prediction are double hybrids, led by PWPB95-D3 [208] (3.61% regularized RMSE) and DSD-PBEPBE-D3 [209] not too far behind (3.76% regularized RMSE). B2PLYP is the only double hybrid functional tested to have a regularized RMSE in excess of 5%, and seven out of the remaining eight yield regularized RMSE around 4.5% or less. This indicates that most double hybrid functionals tested have performance comparable to CCSD for predicting dipole moments, and the top two appear to be better overall for our dataset. This high level of accuracy is somewhat surprising as all the double hybrid calculations at equilibrium geometry were done using restricted orbitals, and thus poor performance appeared nearly inevitable for spin-polarized systems. The errors over the SP subset are indeed substantially larger than the errors over the NSP subset (being nearly twice as large for quite a few functionals), but are still either less than or comparable to corresponding errors for hybrid functionals using unrestricted orbitals (and are significantly less than the corresponding RMP2 errors), resulting in better performance overall. Many of the double hybrids also have very low RMSE over the NSP subset (2-3%, often substantially better than CCSD), which helps drive down the cumulative error significantly as well. Overall, double hybrid functionals give across the board good performance despite being constrained to use restricted orbitals, and should always be employed to estimate dipole moments of spin unpolarized systems where computationally feasible. They are also very likely to give quite good predictions for spin polarized systems when restricted orbitals are employed to prevent N-representability violations, though there exist a few cases like singlet CH_2 where this approach does not yield satisfactory results.

Hybrid functionals also perform quite well in predicting dipole moments relative to the benchmark, with SOGGA11-X [162] being the best (4.84% RMSE), and no fewer than 22 functionals with less than 6% RMSE. Such functionals are thus quite competitive with CCSD (3.95% RMSE) and double hybrids. This close relative performance of many functionals at hybrid and double hybrid levels indicates that good performance for dipole moments is possible from many functionals. It is interesting that the best hybrid meta-GGAs (hmGGAs) do not substantially outperform the best hybrid GGAs (hGGAs). This suggests that the additional flexibility in current hmGGAs has not necessarily resulted in substantial improvements in densities, which is perhaps not all that surprising in light of the vastness of the meta-GGA functional space [63]. Recently developed hybrid meta-GGAs such as ωB97M-V and SCAN0 are in the top cohort, though they are not statistically the best. Furthermore, most of the best-performing hybrids exhibit RMSEs that are between 30% and 100% larger for the SP cases than the NSP cases. Indeed apart from the top two, the next 10 best-performing functionals would reorder considerably if one considered only the NSP cases.

The best local functionals, mBEEF [182] and SCAN [62], are significantly less accurate than the best hybrids, with RMS errors of 7.56% and 8.44% over the whole database. Local functionals typically underestimate dipole moment magnitudes, while HF and functionals with 100% HF exchange tend to overestimate. It is interesting that the best-performing local functionals are all quite recently developed functionals – the N12 GGA, and the mBEEF, SCAN and MS2 mGGAs. Other recently developed semi-empirical functionals, such as B97M-V and MN15-L perform more poorly for dipole moments, so this provides some evidence that fitting to relative energies can lead to loss of accuracy in other properties. For the NSP subset, the best GGA (N12) outperforms the best mGGA (mBEEF), but this result reverses for the more challenging SP subset, as well as the total statistics. As in hybrids, the gap in performance between the SP and NSP subsets for the best functionals is generally smaller than for the wavefunction-based MP2 method, and is more comparable to CCSD. The statistics for the SP subset (both for locals and hybrids) are quite heavily influenced by outliers, such as those discussed below.

A list of the 11 most challenging species is given in Table 4.2. Somewhat unsurprisingly, 8 of the 11 are from the SP subset, which is a consequence of the best functionals being double hybrids that we have constrained to operate on restricted references alone. Perhaps the most challenging species is NaLi. WFT methods all agree that the polarity is $\mathrm{Na^+Li^-}$ and the CCSD(T) $\mu = 0.48$ D is close to the experimental value of 0.46 D [210]. Many functionals (both locals like BLYP [58, 151] and B97M-V [63] and hybrids like MN15 [126] and BMK [92]) however predict the opposite polarity without any apparent systematic reason. Furthermore, BMK [92], MN12-L [125], MN12-SX [125], M11 [123], M06-HF [121] and SOGGA11 [161] predict errors in excess of 100%, with BMK [92] dropping from being one of the very best (4.41% RMS error otherwise) to one of the worst due to this one data-point, and MN12-L [125] predicting an absurd dipole moment of 6.94 D! Fortunately, some functionals do well: ωB97X-V [108], ωB97X-rV [108, 164], ωB97X-D3 [167] and PWB6K [180] predict errors around 1% or less.

The three challenging NSP species are BF and triple bond containing HCCF and HC-

NSP	**SP**
BF	H_2O-F
HCCF	NaLi
HCCCl	H_2O-Cl
	H_2O-Li
	CF
	BS
	CF_2
	NOCl

Table 4.2: The 11 most difficult species in the dataset. These were selected on the basis of the first quartile of absolute DFA errors for each species, with the selected species giving $\geq 5.5\%$ first quartile error. This corresponded to a break in the distribution. Coincidentally, this is also close to the first quartile of RMSEs for functionals considered (5.64%).

CCl. There are some striking systematics in functional performance in these challenging spin unpolarized cases. 86 out of 88 functionals overestimate μ(BF), 82 of 88 underestimate μ(HCCF) and 84 of 88 underestimate μ(HCCCl). These systematics suggest that a hitherto unrecognized delocalization error may be at play here, potentially connected to the $C{\equiv}C{-}X$ moiety. Other spin unpolarized boron containing species like CH_3BH_2 and BH_2F are problematic for functionals from the first four rungs of Jacob's ladder (having first quartile errors $> 5.5\%$ when double hybrids are excluded), which also hint at potential delocalization errors involving bonds to boron. Surprisingly, the strongly correlated O_3 molecule gives relatively little trouble, although most functionals are forced to break spin-symmetry.

The performance for open-shell and incomplete valence species is generally not significantly poorer than for closed shell species for most functionals from Rungs 1-4 of Jacob's ladder, although some radical species are challenging for all functionals, notably the complexes of Li, F and Cl with H_2O (the equivalent Al complex however is surprisingly well behaved), and some small fluorine-containing radicals, such as CF_2 and CF. HF theory predicts incorrect polarity for some species like CO, NF and NO, causing functionals with large amounts of HF exchange (like M06-HF [121] or 50% HF exchange containing PBE50) to also reverse the dipole direction. The opposite problem occurs for some cases like OF, where HF theory gets the direction right but overestimates the magnitude significantly, while most local functionals like PBE [57] and BLYP [58, 151] incorrectly obtain the reverse polarity.

4.3.3 Basis Set Convergence

HF dipole moments converge quite rapidly with increasing basis set size, with the regularized RMS deviation between aug-cc-pCVQZ and aug-cc-pCV5Z values being 0.04% (with a maximum regularized deviation of 0.19%). Even aug-cc-pCVTZ μ have a regularized RMS deviation of only 0.4% relative to aug-cc-pCV5Z numbers (with a maximum regularized deviation of 1%) indicating that the aug-cc-pCV5Z values are essentially at the CBS limit.

Most of the DFAs from Rungs 1-4 of Jacob's ladder exhibit similarly rapid convergence, with a regularized RMS difference of $\leq 0.1\%$ between aug-pc-4 and aug-pc-3, and $\approx 1\%$ or so between aug-pc-4 and aug-pc-2 (a full list of regularized RMS differences is provided in the Supporting Information). Expensive aug-pc-4 calculations of dipole moments are thus unnecessary for practical purposes, as aug-pc-3 and even aug-pc-2 numbers are likely to be sufficiently close for most functionals (especially considering that these functionals typically deviate from benchmark by $> 5-6\%$).

Functional	RMSE		ME		MAX	
	aug-pc-3	aug-pc-2	aug-pc-3	aug-pc-2	aug-pc-3	aug-pc-2
M08-HX	1.06	2.29	0.04	0.5	10.54	10.13
BMK	1.09	1.79	0.02	0.17	9.62	17.72
M06	1.3	3.6	0.01	-0.3	12.77	40.75
M11	1.39	2.4	0.03	0.73	7.54	13.1
M06-HF	2.47	4.12	-0.68	0.21	13.07	39.5
MN15-L	2.68	1.61	0.06	0.26	29.63	14.02
M06-L	2.93	4.58	-0.31	-0.95	27.82	48.77
M11-L	3.94	8.32	-0.19	0.06	34.18	83.57
MN12-SX	5.4	9.08	-0.3	-0.6	62.96	78.07
SOGGA11	11.51	13.55	1.97	2.4	37.89	40.1
MN12-L	18.39	7.86	1.21	0.13	223.74	91.5

Table 4.3: Percentage regularized differences between aug-pc-4 basis μ and aug-pc-3/aug-pc-2 μ for DFT Functionals with $> 1\%$ regularized RMS difference between aug-pc-4 and aug-pc-3 values. A full table with deviations for all functionals is supplied in the Supporting Information.

A few functionals from Rungs 1-4 are however exceptional in having a 1% or larger difference between aug-pc-3 and aug-pc-4 values, suggesting that their density has potentially not yet fully converged even at the quintuple ζ aug-pc-4 level. A full list of these potentially problematic functionals is given in Table 4.3, along with the mean and maximum deviations by magnitude. The poor basis set convergence of intermolecular interaction energies for several of the listed functionals has been noted in the past [211]. It highlights the need to exercise caution when these functionals are being used to estimate molecular properties, due to the lack of an easily approachable convergence limit. SOGGA11 in particular has an inaccessible CBS limit, with 99 species out of 152 having a regularized difference of $> 5\%$ between aug-pc-3 and aug-pc-4 μ, resulting in a very high 11% regularized RMS difference between the two basis sets that is not outlier driven (unlike, say the case of MN12-L—where only 6 species have a $> 5\%$ deviation). It thus appears that the electronic density for SOGGA has likely not converged sufficiently even when the very large aug-pc-4 basis is employed.

Double hybrid functionals exhibit relatively slower basis set convergence compared to other DFAs on account of their dependence on virtual orbitals. The SCF component of μ for all the tested double hybrids however converge at a rate similar to HF and most other functionals ($0.05-0.25\%$ regularized RMS deviation between aug-cc-pCV5Z and aug-cc-pCVQZ, and $< 0.5\%$ regularized RMS difference between aug-cc-pCV5Z and aug-cc-pCVTZ), and aug-cc-pCV5Z numbers are thus expected to be very close to the CBS limit. The $3.5-5\%$ RMS error associated with μ estimated from double hybrid functionals suggest that the non MP2 component of aug-cc-pCVQZ results are probably adequately close to be CBS for practical purposes. The MP2 component has slower basis set convergence, but the behavior can be well approximated by the extrapolation formula $\mu_n^{\text{corr}} = \mu_\infty^{\text{corr}} + A/n^3$ [202]. We therefore recommend that the CBS MP2 component be estimated via extrapolation from aug-cc-pCVQZ and aug-cc-pCVTZ numbers, and be combined with the SCF component at aug-cc-pCVQZ level to obtain a practically useful estimate for μ that does not require any expensive aug-cc-pCV5Z calculations.

4.4 Dipole Moments at Non-equilibrium Configurations

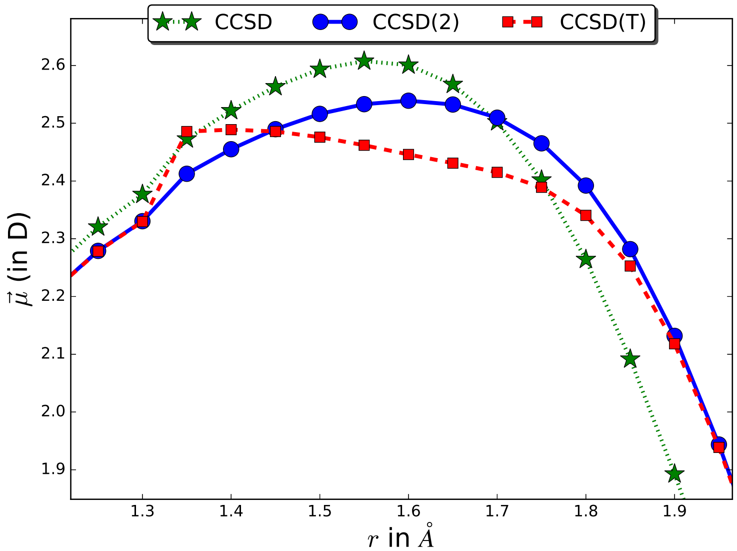


Figure 4.2: Performance of CC methods in predicting $\vec{\mu}$ near Coulson-Fischer point ($r_e = 1.35$ Å) of FH. Note the pronounced kink in the CCSD(T) curve. r is the internuclear separation.

Exploration of non-equilibrium behavior by examining $\vec{\mu}$ along the PES of FH and FCl reveals serious qualitative issues with many of the methods discussed so far. Just as the energy for bond-stretched configurations is a challenge for electronic structure theory, so too is the dipole moment, because such configurations reflect the effect of stronger electron correlations on the property of interest. Indeed, the accuracy of CCSD(T) itself becomes questionable close to the Coulson-Fischer [212] point of the FH PES due to the appearance of kinks [203]. This led us to investigate CCSD(2) [207] as an alternative benchmark in such situations, due to its renormalization of the one-body terms. [213] The performance of CCSD, CCSD(T) and CCSD(2) around the Coulson-Fischer point of FH dissociation (1.35 Å internuclear separation) is given in Fig 4.2. CCSD and CCSD(2) give reasonable behavior, but CCSD(T) has a kink at the spin polarization transition, indicating that CCSD(2) is more reliable. The two methods however agree at shorter and longer distances, and no other unreasonable behavior is observed. No similar issues arise in the case of the FCl molecule, where CCSD(T) and CCSD(2) yield dipole moment curves that do not demonstrate any unphysical behavior and are in perfect visual agreement. Thus CCSD(T) $\vec{\mu}$ for FCl dissociation can be treated as a benchmark.

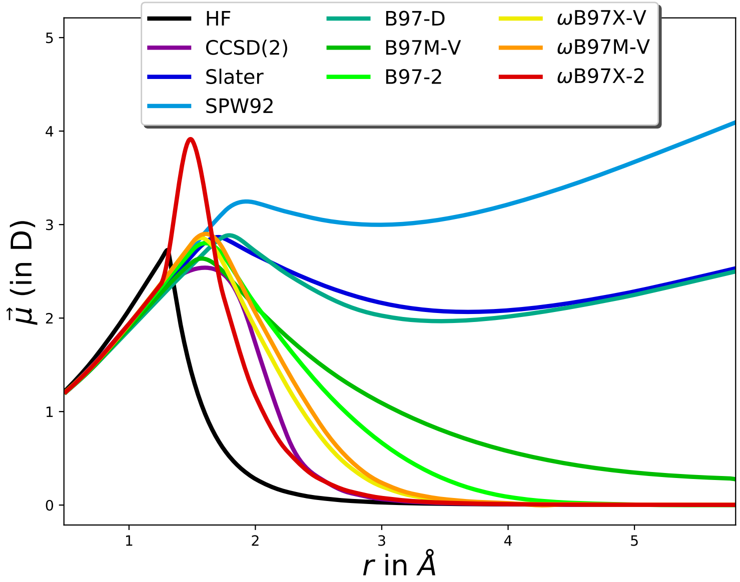


Figure 4.3: Performance of selected DFAs in predicting $\vec{\mu}$ for FH dissociation.

The performance of DFAs in predicting $\vec{\mu}$ along the FH PES varies greatly with internuclear separation, r, as shown in Fig 4.3. The behavior at $r \leq r_e$ is qualitatively fine for all methods, but the performance around the Coulson-Fischer point and larger r is much poorer. Local functionals often spuriously predict a residual charge on the atoms even at very large r [78, 79], causing μ to incorrectly grow asymptotically with r. Surprisingly, SPW92/SVWN [54, 149] leaves behind a larger residual atomic charge than pure Slater [138] exchange without correlation, as can be seen from relative slopes in Fig 4.3. GGAs in general are able to partially reverse this, returning to pure Slater like behavior.

Moving higher up Jacob's ladder to mGGAs alleviates this issue somewhat further, though there are distinct differences between functionals. TPSS [61], revTPSS [153] and τ-HCTH [181] act similarly to GGAs, but SCAN [62], MS2 [172], mBEEF [182], B97M-V [63] and B97M-rV [164] have a smaller residual atomic charge by nearly an order of magnitude (this is not apparent in Fig 4.3 as the B97M-V asymptotic regime is not yet reached). Rather surprisingly, MS1 [172], MVS [173], revM06-L [183], M06-L [101], M11-L [124], MN12-L [125] and MN15-L [127] appear to leave behind no partial charge at all.

On the other hand, hybrid functionals are universally able to eliminate this delocalization issue, allowing μ to go to the correct asymptotic limit of zero. The overall decay rate is typically slower than that of the CCSD(2) benchmark (except for functionals with 100% HF exchange like HFLYP [171], which tend to decay *too* fast like HF), indicating that the asymptotic tail of the μ curve is problematic for modern DFAs. Hybrids also tend to overestimate the maximum μ considerably (as can be seen in Figure 4.3). These difficulties relate to strong electron correlation effects, and represent challenges for future DFA development. Nonetheless, it is encouraging the functionals that are most accurate for chemistry, such as ωB97M-V [96], are amongst the most accurate relative to the CCSD(2) benchmark.

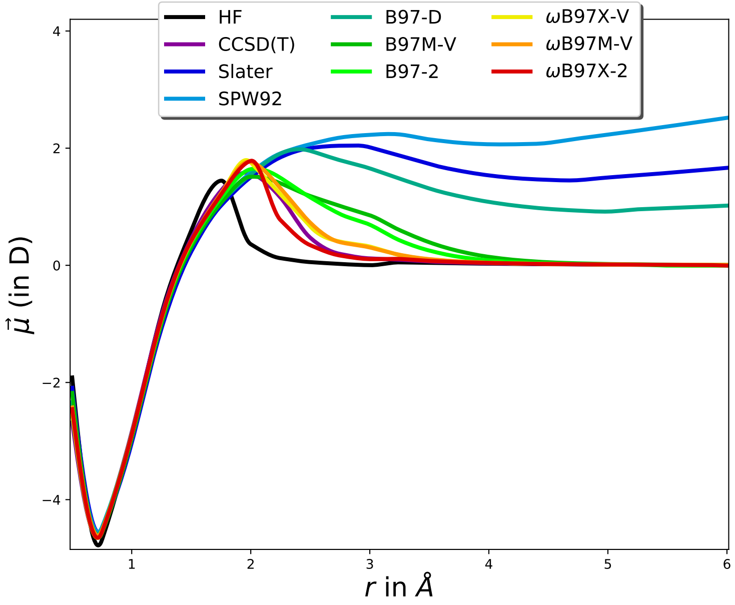


Figure 4.4: Performance of selected DFAs in predicting $\vec{\mu}$ for FCl dissociation.

Rung 5 DFAs however do not fare as well in the stretched bond regime. These functionals inherit N-representability violations present in unrestricted MP2, and consequently may exhibit an unphysical spike in the value of $\vec{\mu}$ around the Coulson-Fischer point (as can be seen in Fig 4.3). Using restricted orbitals (as in the equilibrium case) however yielded asymptotically divergent $\vec{\mu}$ in the bond dissociation limit, as these orbitals artificially enhance the ionic character in stretched bonds. Consequently, $\vec{\mu}$ calculated from unrestricted orbitals is the better (albeit flawed) choice for highly stretched bonds, as it correctly decays to 0 as $r \to \infty$. The decay rate appears to be too large relative to reference values, which is likely a consequence of the larger amount of HF exchange typically employed by double hybrids.

As shown in Fig 4.4, the FCl dissociation is less problematic for functionals, despite the larger median DFA error of 6.65% at equilibrium. LDAs and GGAs still predict a residual charge on the atoms, but the magnitude is significantly lower due to the smaller electronegativity difference between F and Cl (versus F and H). On the other hand, all mGGAs except the aforementioned problematic cases of TPSS, revTPSS and τ-HCTH (which give GGA like behavior) are able to correctly predict neutral atoms at dissociation, showing that the extra physics in the mGGAs is in principle capable of eliminating a weak manifestation of the delocalization error. In fact, mGGAs like B97M-V match the performance of hybrids like B97-2 in this case, although the best performers are still hybrids like ωB97X-V. None of the double hybrids here displayed a sharp spike in μ, indicating that N-representability was likely not violated in this case despite use of unrestricted orbitals.

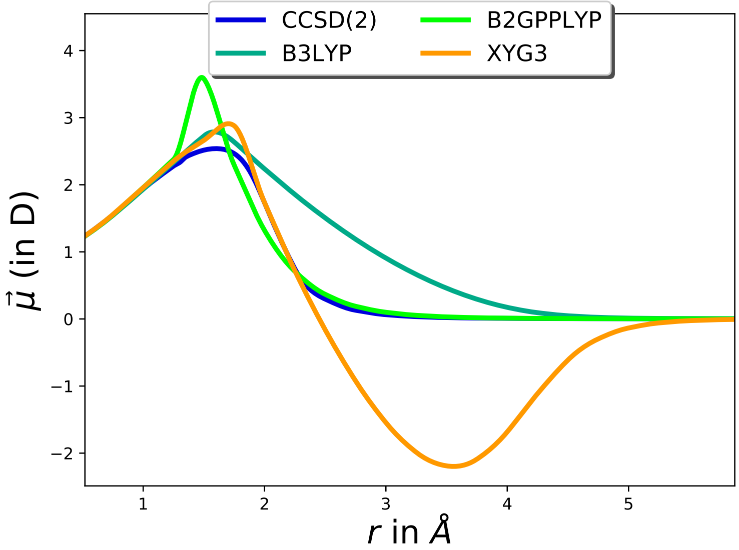


Figure 4.5: $\vec{\mu}$ predicted by the XYG3 functional for FH dissociation, against the CCSD(2) benchmark and the related B3LYP and B2GPPLYP functionals.

One interesting phenomenon unique to the XYG3 [188] and XYGJ-OS [189] double hybrid functionals was observed for both the FH and FCl dissociation curves. These two functionals spuriously reversed the polarity of $\vec{\mu}$ by a considerable amount at moderately large r, as demonstrated in Fig 4.5 for the case of FH dissociation with the XYG3 functional. The origins of this unphysical behavior and any potential implications it may have for development of double hybrid functionals are reported in Ref 81.

4.5 Conclusion

In summary, we find that density functionals are quite good at predicting μ, with the best DFAs having an RMS error of 3.5-4.5%, which is not particularly different from the CCSD RMS error of 3.95%, at a significantly lower computational cost. Most of this success is on account of excellent performance by double hybrid functionals, which yield performance similar to or even slightly superior than CCSD for most species. Hybrid functionals also perform quite well, with 22 functionals having RMS error $< 6\%$, which is only about 50%

worse than CCSD at an even lower computational cost. This success story however hides the fact that many of the best performing hybrid functionals like PBE0 and B97-2 are quite old, and newer functionals like SOGGA11-X [162] and ωB97X-V [108] are only comparable in accuracy. Furthermore, hybrid mGGAs give no improvement in performance compared to the best hybrid GGAs, despite their more sophisticated form. This may reflect the vastness of the mGGA functional space and the previous emphasis on improving energetics (where the best hybrid mGGAs are superior to the best hybrid GGAs).

Assuming that errors in $\vec{\mu}$ are an acceptable proxy for errors in electron density, we conclude that aside from the relatively recent class of double hybrid functionals, there has not been a great deal of progress in improving electron densities in recent years, which is consistent with other recent studies [115, 131]. While the best performing GGA (N12) and mGGA (mBEEF) are relatively recent and are considerably superior to classics like PBE or TPSS, there also exist a number of recently developed functionals that perform comparable to or worse than LSDA, indicating very uneven progress along the axis of improving densities.

On the other hand, one encouraging note is the fact that the mBEEF mGGA is in fact the *best* semi-local functional tested (with SCAN not being very far behind, despite not being fitted to bonded systems), so there are grounds for optimism that future hybrid mGGAs can also do better for predicting μ than the best present hybrids. The ability of some mGGAs to eliminate spurious partial charges during dissociation of FH and FCl is another hopeful feature. Thus future DFA development, particularly involving mGGAs and hybrid mGGAs, could likely benefit from fitting to, or testing on $\vec{\mu}$ in order to seek further improvements, and avoid major failures (e.g. the case of MN12-L for NaLi). We hope that our dataset of computed benchmark dipole moments will assist in development and/or testing of future functionals that improve energetics and also predict highly accurate densities (by contrast, direct comparison against experimental μ values folds in errors due to non-electronic effects such as vibrational averaging). This would take us closer to the exact functional, and in the process assist scientists who are investigating systems where energetics alone are not the sole property of interest. It may likewise be valuable to develop reliable benchmark databases of computed responses of the electron density to perturbations, such as polarizabilities. Previous studies have assessed the performance of DFAs in reproducing experimental polarizabilities [147, 214], but the presence of nuclear quantum effects necessitates the development of a database of benchmark equilibrium polarizabilities, which we present in Chapter 6.

4.6 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: geometries, benchmark dipole moments, table of errors and basis set convergence.

Chapter 5

Second Cumulants of the Electron Density

5.1 Introduction

Quantum chemistry methods have seen increasingly widespread use over the last few decades, and now permit sufficiently accurate computation of energies ranging in scale from total atomization energies [56] to weak noncovalent interactions [215]. In particular, density functional theory [55, 90, 137] (DFT) is ubiquitously employed to study medium to large systems as it provides an acceptable balance between accuracy and computational cost. Although exact in theory [41, 46], practical usage of DFT almost always entails use of computationally tractable density functional approximations (DFAs) within the Kohn-Sham (KS) formalism [48]. It is often remarked that DFAs are not really systematically improvable the way the exact coupled cluster (CC) hierarchy is, but the best-performing DFAs on a given rung of Jacob's ladder [99] statistically improve upon lower rungs [55, 98, 111, 112], indicating that greater complexity *could* (but is not guaranteed to) lead to better accuracy. Caution is still warranted while applying KS-DFT to systems with substantial levels of delocalization error [64, 83, 95, 216] or multireference character [109, 217, 218], but DFAs are quite effective overall at predicting ground state relative energies associated with main-group chemistry [55, 111], and even organometallic chemistry. [219, 220] Indeed, modern DFAs often surpass the accuracy of single-reference wavefunction approaches like second-order Møller-Plesset perturbation theory (MP2) or CC singles and doubles (CCSD) for such problems [80, 98, 221].

This is perhaps unsurprising, as DFA development has historically emphasized improved prediction of such chemically relevant energies, out of the belief that this leads to better approximations to the exact functional. However, the exact functional ought to map the exact density to the exact energy, and it is entirely possible for a DFA to obtain a reasonable energy from a relatively inaccurate density via hidden cancellation of errors. It has indeed been suggested that modern DFAs predict worse densities than older, less parameterized models [115]-leading to considerable discussion [50, 80, 116, 128–131, 133–135, 142] in the scientific community about density predictions from DFT. It is certainly plausible that greater complexity in the functional form creates additional avenues for incorrect behavior (as can be seen in wavefunction theory as well [222, 223]), even without accounting for overfitting to empirical data.

Separate from the debate about "true paths" for DFA development, [115] we note that densities are important for a very practical reason. The electron density controls the response of the electronic energy to external electric fields, such as those that might be encountered in condensed phases or in spectroscopic simulations. Density errors are thus related to energy errors under a different one-particle potential (that DFT should still be formally capable of addressing [41]). Application of DFT beyond ground state gas-phase chemical physics is thus reliant upon DFT yielding reasonable densities, or at least a characterization of when and how various DFAs fail. This has led to investigations on the performance of DFT for predicting dipole moments [80, 81, 141, 147, 224, 225] and polarizabilities [147, 214, 225–228], as they represent the linear and quadratic responses, respectively, of the energy to a spatially uniform electric field.

The natural next step would be to assess the linear-response of the energy to a (spatially) constant electric field gradient, which is equivalent to the molecular quadrupole moment $\mathbf{Q}$. Indeed, there have been a number of assessments of the quality of various density functionals for evaluating molecular quadrupole moments in the past [229–232], albeit before the development of many modern functionals. At the same time, efforts have been made to converge high-quality coupled cluster theory calculations of quadrupole moments with respect to basis set, [233–235] setting the stage for benchmark assessments. $\mathbf{Q}$ is however not translationally invariant for polar systems, making it a somewhat unsuitable metric. We thus decided to investigate the performance of DFT in predicting the second spatial cumulant of the electron density ($\boldsymbol{\mathcal{K}}$) instead, which is a translationally invariant quantity connected to $\mathbf{Q}$ and the dipole moment $\boldsymbol{\mu}$. $\boldsymbol{\mathcal{K}}$ thus relates the density to the response of the energy to electric fields that linearly vary with spatial coordinates. Mathematically, $\boldsymbol{\mathcal{K}}$ is also the variance of the electron density. To the best of our knowledge, there have been no previous studies that explicitly focus on prediction of $\boldsymbol{\mathcal{K}}$ by DFT or MP/CC methods. However, a related quantity had previously been studied with Hartree-Fock (HF) [236] and found to be useful in predicting steric effects of substituents.

5.2 Second Spatial Cumulant of Electron Density

5.2.1 Definition

Let us consider a system with electron number density $\rho(\mathbf{r})$ and total number of electrons $m = \int \rho(\mathbf{r}) d\mathbf{r}$. The spatial probability density $p(\mathbf{r})$ for a single electron is:

$$p(\mathbf{r}) = \frac{\rho(\mathbf{r})}{m} \tag{5.1}$$

An alternative definition of $p(\mathbf{r})$ in terms of the wavefunction $|\Psi\rangle$ is:

$$p(\mathbf{r}) = \langle \mathbf{r}| \left(\underset{m-1}{\mathbf{Tr}} \left[|\Psi\rangle \langle\Psi| \right] \right) |\mathbf{r}\rangle = \frac{\Gamma_1(\mathbf{r},\mathbf{r})}{m} \tag{5.2}$$

which corresponds to the scaled diagonal elements of the one particle density operator, $\Gamma_1(\mathbf{r},\mathbf{r}')$ that results from tracing out $m-1$ degrees of freedom from the full density operator $\Gamma_m = |\Psi\rangle \langle\Psi|$. For a single Slater determinant, this is equivalent to averaging the square of each occupied spin orbital.

The first and second spatial moments of the electron density are consequently:

$$\langle \mathbf{r} \rangle = \int \mathbf{r} p(\mathbf{r}) d\mathbf{r} \tag{5.3}$$

$$\langle \mathbf{r}\mathbf{r}^{\mathbf{T}} \rangle = \int \left(\mathbf{r}\mathbf{r}^{\mathbf{T}}\right) p(\mathbf{r}) d\mathbf{r} \tag{5.4}$$

The second spatial cumulant of the electron density is thus:

$$\boldsymbol{\mathcal{K}} = \langle (\delta\mathbf{r}) \left(\delta\mathbf{r}\right)^{\mathbf{T}} \rangle = \langle \mathbf{r}\mathbf{r}^{\mathbf{T}} \rangle - \langle \mathbf{r} \rangle \langle \mathbf{r} \rangle^{\mathbf{T}} \tag{5.5}$$

which is equivalent to the spatial (co)variance of the probability distribution $p(\mathbf{r})$. The individual components are thus of the form:

$$\mathcal{K}_{xx} = \langle x^2 \rangle - \langle x \rangle^2 = \int x^2 p(\mathbf{r}) d\mathbf{r} - \left(\int x p(\mathbf{r}) d\mathbf{r} \right)^2 \tag{5.6}$$

$$\mathcal{K}_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle = \int xy p(\mathbf{r}) d\mathbf{r} - \left(\int x p(\mathbf{r}) d\mathbf{r} \right) \left(\int y p(\mathbf{r}) d\mathbf{r} \right) \tag{5.7}$$

etc., in terms of the individual cartesian directions x, y, z. $\boldsymbol{\mathcal{K}}$ is thus translationally invariant, and represents the "width" (or "spread") of the electron density. We also note that the eigenvalues of $\boldsymbol{\mathcal{K}}$ are rotationally invariant. In addition, while we have only focused on the second-order cumulant in this work, higher-order cumulants of $p(r)$ can similarly be readily defined and utilized, for problems where they might be relevant.

5.2.2 Connection to Multipole Moments

If the system has nuclear charges $\{Z_A\}$ at positions $\{\mathbf{R_A}\}$, the dipole moment $\boldsymbol{\mu}$ and (cartesian/non-traceless) quadrupole moment $\mathbf{Q}$ are:

$$\boldsymbol{\mu} = \sum_A Z_A \mathbf{R_A} - \int \mathbf{r} \rho(\mathbf{r}) d\mathbf{r} = \sum_A Z_A \mathbf{R_A} - m \langle \mathbf{r} \rangle \tag{5.8}$$

$$\mathbf{Q} = \sum_A Z_A \left(\mathbf{R_A}\right) \left(\mathbf{R_A}\right)^{\mathbf{T}} - \int \left(\mathbf{r}\mathbf{r}^{\mathbf{T}}\right) \rho(\mathbf{r}) d\mathbf{r} = \sum_A Z_A \left(\mathbf{R_A}\right) \left(\mathbf{R_A}\right)^{\mathbf{T}} - m \langle \mathbf{r}\mathbf{r}^{\mathbf{T}} \rangle \tag{5.9}$$

Therefore the second cumulant can be expressed as:

$$\boldsymbol{\mathcal{K}} = \frac{1}{m}\left(\sum_{A} Z_A\left(\mathbf{R_A}\right)\left(\mathbf{R_A}\right)^{\mathbf{T}} - \mathbf{Q}\right) - \frac{1}{m^2}\left(\sum_{A} Z_A\mathbf{R_A} - \boldsymbol{\mu}\right)\left(\sum_{A} Z_A\mathbf{R_A} - \boldsymbol{\mu}\right)^{\mathbf{T}} \quad (5.10)$$

5.2.3 Physical Interpretation

Let us begin with the case of atoms, and suppose that the single electron probability distribution for a (spherical) atom at $(0,0,0)$ is $p_1(\mathbf{r})$. Spherical symmetry indicates that:

$$\int xp_1(\mathbf{r})d\mathbf{r} = \int yp_1(\mathbf{r})d\mathbf{r} = \int zp_1(\mathbf{r})d\mathbf{r} = 0 \quad (5.11)$$

$$\int xyp_1(\mathbf{r})d\mathbf{r} = \int yzp_1(\mathbf{r})d\mathbf{r} = \int xzp_1(\mathbf{r})d\mathbf{r} = 0 \quad (5.12)$$

$$\int x^2p_1(\mathbf{r})d\mathbf{r} = \int y^2p_1(\mathbf{r})d\mathbf{r} = \int z^2p_1(\mathbf{r})d\mathbf{r} = \eta^2 \quad (5.13)$$

This implies that $\mathcal{K}_{xx} = \mathcal{K}_{yy} = \mathcal{K}_{zz} = \eta^2$. Therefore $\sqrt{3}\eta$ is an effective atomic radius associated with the spatial extent of the electron density, much like a covalent or van der Waals radius. In particular for the H atom, η is the Bohr radius (1 a.u.=0.529 Å).

For molecules without symmetry, $\boldsymbol{\mathcal{K}}$ is not generally a diagonal 3×3 matrix, and likewise $\mathcal{K}_{xx} \neq \mathcal{K}_{yy} \neq \mathcal{K}_{zz}$. $\boldsymbol{\mathcal{K}}$ can be diagonalized, to yield principal axes, and 3 principal components, very analogous to the molecular inertia tensor associated with the nuclei, $\mathbf{I} = \sum_{A} m_A\left(\mathbf{R_A}\right)\left(\mathbf{R_A}\right)^{\mathbf{T}}$. The eigenvalues of $\boldsymbol{\mathcal{K}}$ thereby give the square of the spatial extent of the electron density along each principal axis. Systems where both a C_n $(n > 1)$ axis of rotation and σ_v plane(s) of symmetry are present (i.e. C_{nv} or higher symmetry) have the principal axes be defined by molecular symmetry, and we shall later choose members of our data set based on this convenient simplification.

5.2.4 Behavior vs System Size

The behavior of $\boldsymbol{\mathcal{K}}$ vs system size could be useful in characterizing the utility of this quantity. A simple case to consider is a 1D lattice of N noninteracting He atoms, placed at $(0,0,0),(0,0,a),(0,0,2a)\ldots(0,0,(N-1)a)$. Let each atom have spatial extent η as in Eq. (5.13).

Subsequently, $p(\mathbf{r})$ for the supersystem is:

$$p(\mathbf{r}) = \frac{1}{N}\sum_{n=0}^{N-1} p_1(\mathbf{r} - na\hat{\mathbf{z}}) \quad (5.14)$$

which is the average of the individual single electron probability distributions. Transverse to the lattice vector $a\hat{\mathbf{z}}$, we thus have:

$$\langle x^2 \rangle = \frac{1}{N}\sum_{n=0}^{N-1}\int x^2 p_1(\mathbf{r}-na\hat{\mathbf{z}})d\mathbf{r} = \frac{1}{N}\sum_{n=0}^{N-1}\int x_n^2 p_1(\mathbf{r}_n)d\mathbf{r}_n = \eta^2 \tag{5.15}$$

$$\langle x \rangle = \frac{1}{N}\sum_{n=0}^{N-1} x\int p_1(\mathbf{r}-na\hat{\mathbf{z}})d\mathbf{r} = \frac{1}{N}\sum_{n=0}^{N-1}\int x_n p_1(\mathbf{r}_n)d\mathbf{r}_n = 0 \tag{5.16}$$

(where $\mathbf{r}_n = \mathbf{r} - na\hat{\mathbf{z}}$). Therefore $\mathcal{K}_{xx} = \eta^2 = \mathcal{K}_{yy}$ and is invariant vs system size.

Parallel to the lattice vector $a\hat{\mathbf{z}}$ however, we obtain:

$$\begin{aligned}\langle z^2 \rangle &= \frac{1}{N}\sum_{n=0}^{N-1}\int z^2 p_1(\mathbf{r}-na\hat{\mathbf{z}})d\mathbf{r} \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\int (z_n+na)^2 p_1(\mathbf{r}_n)d\mathbf{r}_n \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\left(\int z_n^2 p_1(\mathbf{r}_n)d\mathbf{r}_n + n^2a^2\int p_1(\mathbf{r}_n)d\mathbf{r}_n + 2an\int z_n p_1(\mathbf{r}_n)d\mathbf{r}_n\right) \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\left(\eta^2+n^2a^2\right) = \eta^2 + \frac{(N-1)(2N-1)}{6}a^2 \end{aligned}\tag{5.17}$$

$$\begin{aligned}\langle z \rangle &= \frac{1}{N}\sum_{n=0}^{N-1}\int z p_1(\mathbf{r}-na\hat{\mathbf{z}})d\mathbf{r} \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\int (z_n+na)\, p_1(\mathbf{r}_n)d\mathbf{r}_n \\ &= \frac{1}{N}\sum_{n=0}^{N-1}\left(\int z_n p_1(\mathbf{r}_n)d\mathbf{r}_n + na\int p_1(\mathbf{r}_n)d\mathbf{r}_n\right) \\ &= \frac{1}{N}\sum_{n=0}^{N-1} na = \frac{N-1}{2}a \end{aligned}\tag{5.18}$$

$$\therefore \mathcal{K}_{zz} = \langle z^2\rangle - \langle z\rangle^2 = \eta^2 + \frac{(N^2-1)}{12}a^2 \tag{5.19}$$

$\mathcal{K}_{zz}$ thus grows as $O(N^2)$ vs the system size N. This increase is however entirely due to geometric/structural factors (i.e. only dependent on the lattice spacing a and not on the electronic contribution from $p_1(\mathbf{r})$). Thus differences between $\mathcal{K}_{zz}$ computed by various methods will be independent of N, and will solely be a function of the subsystem densities

$\{p_1(\mathbf{r})\}$. This is not true in the interacting subsystems limit, but the analysis nonetheless reveals a significant contribution to $\boldsymbol{\mathcal{K}}$ from molecular geometry alone.

A similar analysis for 2D (square) and 3D (cubic) lattices shows that $\mathcal{K}_{ii}$ grows as $O(N)$ and $O(N^{\frac{2}{3}})$ respectively vs number of identical noninteracting subsystems N (where $\hat{\mathbf{i}}$ is parallel to the lattice vectors), due to geometric factors, while the electronic contribution intrinsically arising from $p(\mathbf{r})$ remains constant. Consequently, differences in $\boldsymbol{\mathcal{K}}$ should be size-intensive in the non-interacting limit. However, relative error in $\boldsymbol{\mathcal{K}}$ would shrink with increasing N (as the reference value in the denominator would increase). We consequently only consider absolute errors of the form $\boldsymbol{\mathcal{K}} - \boldsymbol{\mathcal{K}}_{ref}$, vs a reference value $\boldsymbol{\mathcal{K}}_{ref}$. This stands in contrast to properties like dipole moments and static polarizabilities, which are properly size-extensive and thus suitable for relative/percentage error based metrics. It is also possible to look at standard deviations (i.e. $\sqrt{\mathcal{K}_{ii}}$) instead of variances, but the geometric factors would prevent the errors from being size-intensive in that case.

NSP			SP	
AlF	$\mathrm{Cl_2}$	Mg	$\mathrm{AlH_2}$	NH
Ar	ClCN	$\mathrm{Mg_2}$	$\mathrm{BH_2}$	$\mathrm{NH_2}$
BF	ClF	$\mathrm{N_2}$	BO	$\mathrm{NO_2}$
$\mathrm{BH_2Cl}$	FCN	$\mathrm{NH_3}$	BS	NP
$\mathrm{BH_2F}$	$\mathrm{H_2}$	$\mathrm{NH_3O}$	Be	Na
$\mathrm{BH_3}$	$\mathrm{H_2O}$	NaCl	BeH	$\mathrm{Na_2}$
$\mathrm{BHF_2}$	HBO	NaH	$\mathrm{C_2H}$	NaLi
$\mathrm{BeH_2}$	HBS	Ne	$\mathrm{CF_2}$	$\mathrm{O_2}$
$\mathrm{C_2H_2}$	HCCCl	$\mathrm{OCl_2}$	$^3\mathrm{CH_2}$	$\mathrm{O_3}$
$\mathrm{C_2H_4}$	HCCF	$\mathrm{PH_3}$	$\mathrm{CH_3}$	$\mathrm{OF_2}$
$\mathrm{CH_2BH}$	HCHO	$\mathrm{PH_3O}$	CN	P
$\mathrm{CH_2BO}$	HCN	$\mathrm{SCl_2}$	$\mathrm{F_2}$	$\mathrm{P_2}$
$\mathrm{CH_3Cl}$	HCl	$\mathrm{SF_2}$	H	PCl
$\mathrm{CH_3F}$	HF	$\mathrm{SH_2}$	$\mathrm{H_2CN}$	PF
$\mathrm{CH_3Li}$	HNC	$\mathrm{SO_2}$	HCHS	PH
$\mathrm{CH_4}$	He	$\mathrm{SiH_3Cl}$	HCP	$\mathrm{PH_2}$
CO	$\mathrm{LiBH_4}$	$\mathrm{SiH_3F}$	Li	$\mathrm{PO_2}$
$\mathrm{CO_2}$	LiCl	$\mathrm{SiH_4}$	N	$\mathrm{S_2}$
CS	LiF	SiO	NCl	SO-trip
CSO	LiH		NF	$\mathrm{SiH_3}$
			$\mathrm{NF_2}$	

Table 5.1: The 100 species in the dataset, sorted by whether they are not spin-polarized (NSP) or spin-polarized (SP).

5.3 Dataset

We have investigated $\boldsymbol{\mathcal{K}}$ for 100 small main-group systems (listed in Table 5.1), for which it was possible to get highly accurate benchmark values with CC singles and doubles with perturbative triples (CCSD(T) [23]) at the complete basis set (CBS) limit. The systems were chosen from the combined datasets considered in Refs 80 and 226 such that $\boldsymbol{\mathcal{K}}$ was diagonal for a fixed coordinate system for all (spatial symmetry preserving) electronic structure methods. Asymmetric systems with nondiagonal $\boldsymbol{\mathcal{K}}$ were not included for simplicity. Linear molecules with a single π electron/hole (such as OH) were also not considered, as real valued orbitals would be incapable of predicting a cylindrically symmetric orbital of π symmetry [237, 238], and would thus spuriously lead to $\mathcal{K}_{xx} \neq \mathcal{K}_{yy}$. Complex valued orbitals would thus be necessary to describe such species with proper symmetry within a DFT framework [237].

The complete dataset consists of 59 not spin-polarized (NSP) and 41 spin-polarized (SP) systems. The NSP vs SP classification was done on the basis of whether the stable HF solution has $\langle S^2 \rangle = 0$ or not. NSP species are thus unambiguously closed-shell, and are thus more likely to be 'easier' for single-reference quantum chemistry methods like KS-DFT or MP2. Indeed, the size of the (T) correction to $\boldsymbol{\mathcal{K}}$ was quite small for nearly all species (NSP or SP, as discussed later in Sec 5.4.1), indicating that the multireference character of the chosen systems (if any) did not strongly influence $\boldsymbol{\mathcal{K}}$ predictions and that CCSD(T) is likely to be an adequately accurate benchmark. Furthermore, none of these systems have pathological, delocalization driven qualitative failures [77–81], making them reasonable choices for understanding the behavior of DFAs in the regime where they are expected to work well.

The error $\epsilon_{i,l}$ in an individual component $\mathcal{K}^l_{ii}$ for a given molecule l (vs a reference value $\mathcal{K}^l_{ref,ii}$) is:

$$\epsilon_{i,l} = \mathcal{K}^l_{ii} - \mathcal{K}^l_{ref,ii} \tag{5.20}$$

The cumulative errors over all molecules and directions are thus:

1. Root mean square error (RMSE): $\sqrt{\dfrac{1}{3L}\sum\limits_{l=1}^{L}\left(\epsilon^2_{x,l}+\epsilon^2_{y,l}+\epsilon^2_{z,l}\right)}$.
2. Mean error (ME): $\dfrac{1}{3L}\sum\limits_{l=1}^{L}\left(\epsilon_{x,l}+\epsilon_{y,l}+\epsilon_{z,l}\right)$.
3. Maximum absolute error (MAX): $\max\left(|\epsilon_{i,l}|\right) \forall i \in \{x, y, z\}$ and $\forall l \in \mathbb{Z}^+, l \leq L$.

All quantities are reported in atomic units (a.u.) unless specified otherwise.

5.3.1 Computational Details

All calculations were performed with the Q-Chem 5 package [191], using fixed geometries obtained from Refs 80 and 226 (also provided in the supporting information). We examined the performance of 47 DFAs spanning all five rungs of Jacob's ladder (as can be seen from Table 5.2), ensuring reasonable representation at each level. Individual DFAs were selected based on (perceived) popularity, recency and performance over other benchmark datasets [55, 80, 98, 111, 226].

$\boldsymbol{\mathcal{K}}$ was obtained from $\mathbf{Q}$ computed for this work and $\boldsymbol{\mu}$ obtained from Ref 80. $\mathbf{Q}$ for self-consistent field (SCF) approaches like HF and non double hybrid DFAs were found analytically via integrating over $\rho(\mathbf{r})$. $\mathbf{Q}$ for other methods (MP2, CC or double hybrid DFT) was found via a central, two point finite difference approach using a constant electric field gradient of 2×10^{-4} a.u. (similar to Ref 80). Finite difference errors thus introduced appear to be quite small, as the RMS deviation between analytic and finite-difference CCSD/aug-cc-pCVTZ [70, 198–200] $\mathcal{K}_{ii}$ is 2.2×10^{-4} a.u.

Comparison between wavefunction theory and various DFAs are only meaningful at the CBS limit, due to different rates of basis set convergence. HF/aug-cc-pCV5Z $\mathcal{K}_{ii}$ were assumed to be at the CBS limit as it only has an RMS deviation of 2.5×10^{-4} a.u vs aug-cc-pCVQZ results (which should be sufficiently small, in light of the empirically observed exponential convergence of HF energies vs cardinal number of the basis set [239, 240]). Similarly, DFT/aug-pc-4 [65–69, 241] values were assumed to be at the CBS limit for functionals from Rungs 1-4 of Jacob's ladder. The virtual orbital dependent correlation contribution $\boldsymbol{\mathcal{K}}^{\text{corr}}$ (in MP2/CC/double hybrids) was extrapolated to the CBS limit via the two point extrapolation formula [202] $\mathcal{K}_{ii}^{\text{corr,n}} = \mathcal{K}_{ii}^{\text{corr,CBS}} + \dfrac{A}{n^3}$ from aug-cc-pCVTZ ($n = 3$) and aug-cc-pCVQZ ($n = 4$) (which is adequate for dipoles [80, 202] and appears to also be adequate for quadrupole moments [233–235]).

Local exchange-correlation integrals for all DFT calculations were computed over a radial grid with 99 points and an angular Lebedev grid with 590 points for all atoms. Nonlocal correlation was evaluated on an SG-1 grid [201]. Unrestricted (U) orbitals were employed for all CC (except Be, where UHF breaks spatial symmetry) and non-double hybrid DFT calculations. MP2 is known to yield non N-representable densities when spin-symmetry breaks [203], leading to poor dipole [80] and polarizability [226] predictions. Consequently, MP2 with both restricted (R) and unrestricted orbitals were carried out, and the restricted variant found to yield a significantly smaller RMSE. Double hybrid calculations were subsequently carried out with only R orbitals. Stability analysis was performed in the absence of fields to ensure the orbitals correspond to a minimum, and the resulting orbitals were employed as initial guesses for finite field calculations (if any). The frozen-core approximation was not employed in any calculation.

Method	Class	RMSE			ME	MAX	Method	Class	RMSE			ME	MAX
		Full	NSP	SP					Full	NSP	SP		
CCSD	WFT	0.003	0.002	0.003	-0.001	0.022	SCAN0 [174]	Rung 4	0.008	0.004	0.011	0.000	0.047
MP2	WFT	0.022	0.004	0.035	0.001	0.312	B97-2 [118]	Rung 4	0.009	0.004	0.013	0.001	0.047
RMP2	WFT	0.010	0.004	0.015	0.002	0.074	TPSSh [154]	Rung 4	0.009	0.004	0.013	0.002	0.053
HF	WFT	0.027	0.015	0.039	0.007	0.301	PBE0 [91]	Rung 4	0.009	0.005	0.014	0.003	0.052
RHF	WFT	0.020	0.015	0.026	0.006	0.091	B97 [175]	Rung 4	0.010	0.006	0.013	0.005	0.049
							HSE-HJS [177]	Rung 4	0.010	0.005	0.014	0.003	0.051
SPW92 [138, 149]	Rung 1	0.022	0.017	0.028	0.011	0.129	SOGGA11-X [162]	Rung 4	0.010	0.005	0.015	0.004	0.057
Slater [138]	Rung 1	0.047	0.038	0.057	0.034	0.275	LRC-ωPBEh [179]	Rung 4	0.011	0.005	0.016	0.002	0.061
							PW6B95 [180]	Rung 4	0.011	0.006	0.016	0.003	0.084
BPBE [57, 58]	Rung 2	0.013	0.008	0.017	0.004	0.062	ωM05-D [168]	Rung 4	0.011	0.007	0.015	0.005	0.064
mPW91 [242]	Rung 2	0.014	0.010	0.019	0.006	0.069	M06 [120]	Rung 4	0.011	0.007	0.016	0.001	0.058
B97-D3 [59]	Rung 2	0.016	0.010	0.022	0.007	0.103	ωB97X-V [108]	Rung 4	0.012	0.008	0.016	0.006	0.066
PBE [57]	Rung 2	0.016	0.012	0.021	0.008	0.079	ωB97X-D [165]	Rung 4	0.012	0.006	0.017	0.003	0.064
N12 [159]	Rung 2	0.020	0.013	0.028	0.006	0.120	M06-2X [120]	Rung 4	0.013	0.006	0.019	0.003	0.104
BLYP [58, 151]	Rung 2	0.021	0.016	0.026	0.012	0.112	HFLYP [151]	Rung 4	0.014	0.013	0.015	-0.004	0.051
SOGGA11 [161]	Rung 2	0.048	0.035	0.062	0.024	0.500	M11 [123]	Rung 4	0.014	0.009	0.019	0.005	0.083
							B3LYP [89]	Rung 4	0.014	0.009	0.019	0.007	0.068
SCAN [62]	Rung 3	0.009	0.005	0.012	0.002	0.052	CAM-B3LYP [94]	Rung 4	0.014	0.010	0.019	0.007	0.070
MS2 [172]	Rung 3	0.010	0.004	0.014	0.000	0.072	M08-HX [122]	Rung 4	0.015	0.009	0.020	0.009	0.079
TPSS [61]	Rung 3	0.010	0.006	0.015	0.003	0.065	ωB97M-V [96]	Rung 4	0.022	0.013	0.030	0.013	0.127
mBEEF [182]	Rung 3	0.011	0.005	0.017	-0.001	0.074	MN15 [126]	Rung 4	0.035	0.013	0.052	0.013	0.231
M06-L [101]	Rung 3	0.014	0.010	0.018	-0.006	0.068							
revM06-L [183]	Rung 3	0.015	0.013	0.018	0.002	0.091	DSD-PBEPBE [209]	Rung 5	0.005	0.003	0.007	0.002	0.032
B97M-V [63]	Rung 3	0.026	0.013	0.037	0.011	0.157	XYGJ-OS [189]	Rung 5	0.006	0.003	0.008	0.002	0.028
M11-L [124]	Rung 3	0.038	0.023	0.052	0.005	0.332	PTPSS [208]	Rung 5	0.006	0.003	0.009	0.002	0.032
MN15-L [127]	Rung 3	0.040	0.024	0.055	0.022	0.278	XYG3 [188]	Rung 5	0.006	0.003	0.009	0.002	0.032
							ωB97M(2) [98]	Rung 5	0.008	0.005	0.010	0.003	0.037
							B2GPPLYP [190]	Rung 5	0.008	0.005	0.011	0.004	0.042
							ωB97X-2(TQZ) [187]	Rung 5	0.008	0.006	0.010	0.005	0.044
							B2PLYP [97]	Rung 5	0.010	0.007	0.013	0.005	0.047

Table 5.2: Errors in $\boldsymbol{\mathcal{K}}$ (in a.u) predicted by various DFAs for the dataset. Positive ME values indicate less compact (i.e. more diffuse) densities, relative to the benchmark.

5.4 Results and Discussions

5.4.1 Full Dataset

The errors in DFA predictions for the full dataset are reported in Table 5.2, along with errors for the wavefunction methods CCSD, MP2 and HF. Considering the wavefunction theories (WFTs) first, we see that CCSD has the lowest RMSE of 0.003 of all the methods considered, and gives fairly similar performance across the NSP and SP subsets. MP2 gives good performance for the NSP dataset, but N-representability failures [203] lead to significantly worse results over the SP subset. Use of R orbitals ameliorates this considerably, with RMP2 having a quite improved RMSE of 0.010 over the full dataset. RMP2 is nonetheless still somewhat challenged by open-shell systems (especially NF and NCl) where the artificial enforcement spin-symmetry appears to be suboptimal. We further note that MP3 is also expected to have the same N-representability failures as MP2 and is thus unlikely to lead to significant improvements over MP2. This in fact has been observed for μ, although alternative orbital choices could lead to better MP performance. [221]

HF performs quite poorly due to lack of correlation, having an RMSE of 0.027 (which can be considered as a ceiling for reasonable DFA performance). Based on the mean error (ME), HF has a tendency to make the variance too large (i.e. HF densities are too diffuse). The SP species are much more challenging than NSP, and RHF actually reduces RMSE substantially to 0.020. However, this is almost solely on the account of two challenging alkali metal containing species (Na_2 and NaLi) and the RMSE for HF and RHF are quite similar upon their exclusion. Nonetheless, this serves as a warning that spin-symmetry breaking might compromise property predictions, despite being the natural approach for improving the energy. In particular, spin-symmetry breaking in diatomics like Na_2 or NaLi results in densities similar to two independent atoms vs a bonded molecule, leading to larger widths than the restricted solution.

Coming to the DFAs, we observe that rung 1 local spin-density approximations (LSDA) fare considerably worse than CCSD or RMP2. Bare Slater exchange has a rather large RMSE of 0.047, while inclusion of correlation reduces RMSE substantially to 0.022 for SPW92. The SP subset is significantly more challenging in both cases, with an RMSE that is almost double the corresponding NSP value. The minimally parameterized nature of the LSDA functionals mean that the SPW92 RMSE of 0.022 can also be considered as a reasonable reference for judging DFAs against. In fact, the HF, RHF and SPW92 RMSEs collectively suggest that DFAs with RMSE larger than 0.02 are not really accurate for predicting $\boldsymbol{\mathcal{K}}$.

Moving up Jacob's ladder to rung 2 generalized gradient approximations (GGAs), we observe that all but one of the functionals investigated improve upon SPW92. SOGGA11 is the exception, being essentially as bad as bare Slater exchange. On the other hand, BPBE and mPW91 have substantially lower RMSEs, indicating considerable improvement over rung 1. However, the SP RMSE continues to be nearly twice the NSP one for all GGAs.

Rung 3 meta-GGAs (mGGAs) further improve upon predictions, with the best mGGA (SCAN) predicting a quite low RMSE of 0.009, while MS2 and TPSS are also fairly reasonable. It is worth noting that all three were developed with a heavy emphasis on non-empirical constraints and were specifically fit to the H atom (one of the most challenging species in the dataset, as discussed later). In contrast, modern, mGGAs developed principally by fitting to empirical benchmark data appear to fare worse, with MN15-L, M11-L and B97M-V being particularly disappointing as they fare worse than SPW92. Nonetheless, it is worth noting that the modern, empirically fitted mBEEF functional yields a respectable performance.

Hybrid functionals on rung 4 of Jacob's ladder do not significantly improve upon mGGAs. The best-performer is SCAN0, which only marginally improves upon the parent SCAN functional. Similarly, TPSSh only slightly improves upon TPSS. This general behavior stands in contrast to the case of dipole moments, where hybrid functionals strongly improve upon lower rungs [80]. Some other decent performers are B97-2 (which was partially fitted to densities) PBE0, B97, HSE-HJS and SOGGA11-X which also happen to be hybrid GGAs. PW6B95 is the best-performing hybrid mGGA that was not already based on an existing Rung 3 functional with good performance. Interestingly, HFLYP improves significantly upon HF, mostly by virtue of dramatically reducing SP errors (in large part because it does not break spin-symmetry for NaLi or Na_2). SP errors are roughly 2-3 times the NSP errors for nearly all other hybrid functionals. It is also worth noting that that modern ωB97M-V and MN15 functionals fare particularly poorly, yielding performance worse than SPW92. On the other hand, several empirically fitted functionals like M06 and ωB97X-V yield quite respectable performance (albeit significantly worse than SCAN0).

Rung 5 double hybrid functionals reduce error significantly vs hybrids, with DSD-PBEPBE approaching CCSD levels of RMSE for NSP species (although SP species are much more challenging). Even B2PLYP (the poorest performer, and coincidentally, the oldest) performs similarly to a good hybrid functional like PBE0. Most rung 5 functionals also improve upon RMP2 (especially for the SP subset), indicating perceptible benefit from the local exchange-correlation contributions. It is also worth noting that the recently developed ωB97M(2) functional yields mediocre performance, despite being one of the best-performers for energy predictions [98] (although it represents an enormous improvement over the parent ωB97M-V functional).

The positive ME values in Table 5.2 also indicates that most DFAs predict slightly less compact densities (especially LSDA). This is perhaps not too surprising in light of delocalization error present in the studied functionals [95]. However, the connection between density compactness and delocalization error is not always straightforward, as shown later. Indeed, it can be seen that the local M06-L and mBEEF functionals systematically predict too small $\boldsymbol{\mathcal{K}}$, which is contrary to expectations based on delocalization error alone. HFLYP however has a negative ME, consistent with overlocalization of density that is expected from 100% HF exchange. On the other hand, HF systematically overestimates $\boldsymbol{\mathcal{K}}$, which seems puzzling at a first glace as HF should overlocalize if anything [88, 95]. However, lack of correlation in HF leads to artificial symmetry breaking where densities become more 'atom-like' than 'bond-like', leading to spurious overestimation of $\boldsymbol{\mathcal{K}}$ for NaLi and Na_2. In fact, lack of correlation hinders electrons from occupying the same region of space and could lead to wider 'spread' in the electron density.

Comparison to the results obtained earlier for dipoles [80] indicates some similarities and differences. Ascending Jacob's ladder leads to improved predictions in both cases, and several modern, empirically designed functionals are found to perform relatively poorly. In fact, some of the functionals that were found to be amongst the top performers for their rung in the dipole study (SCAN, DSD-PBEPBE, PBE0, SOGGA11-X etc.) continue to perform well. The major differences are that hybrid functionals do not significantly improve upon mGGAs and some functionals that are good for $\boldsymbol{\mathcal{K}}$ predictions do not fare as well for dipoles. A clear example of this is SCAN0 faring worse at predicting $\boldsymbol{\mu}$ than ωB97M-V, despite the former being the best hybrid for predicting $\boldsymbol{\mathcal{K}}$ (while the latter is amongst the worst). Nonetheless, the ability of several functionals to predict both $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mu}$ with low error is encouraging, as that indicates reasonable performance for problems involving external electric fields.

It is also worthwhile to identify what species in the dataset are most challenging for DFAs. Ordering the molecules in the dataset by the first quartile of RMS error (over cartesian axes) across different DFAs reveals that there is a break in the distribution after the first

10 species. These difficult cases are: Be, H, Li, NaLi, Na_2, BeH, LiH, H_2, BeH_2 and BF in descending order of difficulty. Several of these cases are also known to be challenging for dipole moment [80] and static polarizability [226] predictions as well. NaLi in particular was already known to be a very challenging case for many DFAs, and it is unsurprising that the analogous Na_2 is also challenging. Both of these alkali metal dimers feature long bonds (~ 3 Å) and are SP at equilibrium geometry, suggesting some multireference character (analogous to the isovalent case of stretched H_2). In addition, Na_2 and NaLi are outliers with respect to RMS deviation of CCSD $\mathcal{K}$ vs the CCSD(T) benchmark, having ~ 0.01 deviation (while the next largest deviation is about half of that, for CH_3Li). Near-exact calculations with the adaptive sampling configuration interaction method [243, 244] (with the small cc-pVDZ basis [200]) however indicate that the species are not particularly multireference (a single determinant has $\sim 90\%$ weight in the total wavefunction) and CCSD(T) is sufficiently accurate for $\mathcal{K}$ (as shown in the supporting information). Furthermore, the three atoms represent the most challenging cases overall, with Li having a first quartile RMSE $\sim 35\%$ larger than NaLi.

5.4.2 The Case of Challenging Atoms

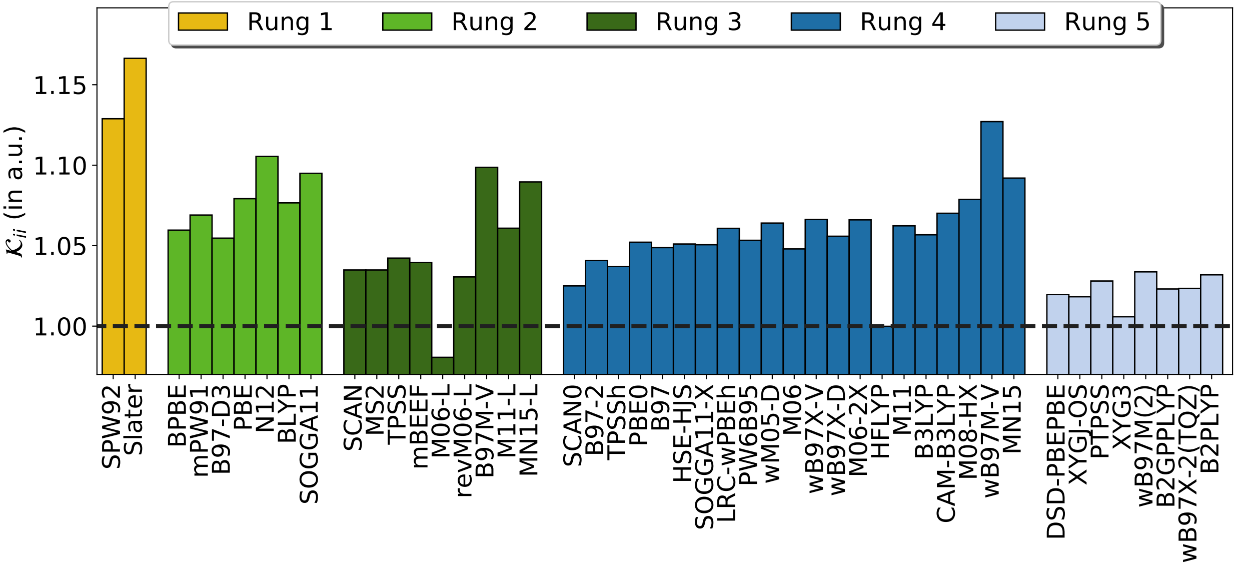


Figure 5.1: $\mathcal{K}_{ii}$ predicted for the H atom, by various DFAs (ordered by their position in Table 5.2). The dashed line is the analytic value.

It is thus worthwhile to examine these atoms in greater detail for better insight into functional behavior. H has only one electron and thus all DFA errors are definitionally self-interaction. It is thus interesting to observe that the error in $\boldsymbol{\mathcal{K}}$ does not appear to correlate well with the delocalization error in many cases, as can be seen from Fig 5.1. Perhaps the most clear example of this is a comparison between the local M06-L functional and the range separated hybrid ωB97M-V, with the former surprisingly predicting a *smaller* H atom than exact quantum mechanics, while the latter overestimates the size to essentially the same extent as SPW92 (contrary to the behavior seen for delocalization error [95]). Similarly, M06-2X predicts a larger H atom than M06 despite having twice the HF exchange (54% vs 27%). It is thus apparent that the local exchange-correlation components of several modern density functionals lead to significant errors for the size of the H atom, which are counterintuitive from the perspective of delocalization error (or fraction of HF exchange present).

On the other hand, if we only consider functionals that use the same local exchange-correlation components hybridized with varying amounts of HF exchange (such as PBE/PBE0), we find that error decreases with increase in the HF exchange contribution. This is on account of the functional becoming strictly more HF like. It is also worth noting that the double hybrid functionals appear to have much lower error than hybrids, potentially due to relatively smaller contributions from local exchange-correlation. While cancellation of errors in larger systems is likely to make this less of an issue (as can be seen from other species in the dataset), future density functional development would likely benefit from including the size and polarizability [226] of the H atom as soft constraints.

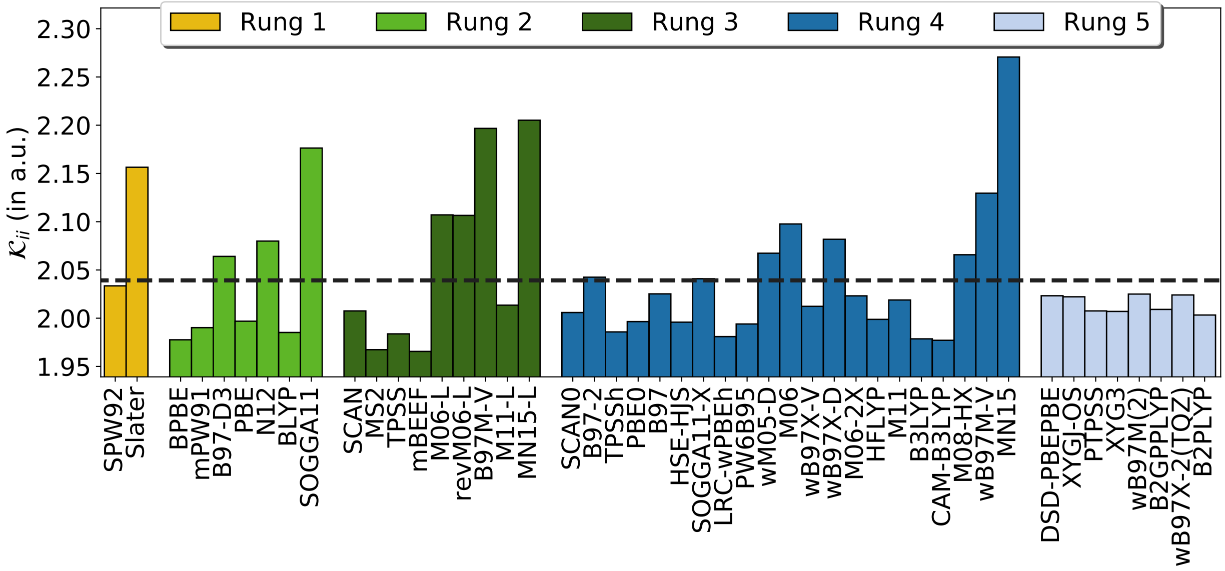


Figure 5.2: $\mathcal{K}_{ii}$ predicted for the Li atom, by various DFAs. The dashed line is the benchmark (CCSD(T)/CBS) value and the HF value is 2.070.

Li superficially resembles H if the core electrons are neglected. However, Fig 5.2 shows that the error $\boldsymbol{\mathcal{K}}$ predictions are quite different. Most functionals *underestimate* the size of the diffuse alkali metal atom, predicting a more compact density than the benchmark. Delocalization errors are normally expected to lead to less compact densities due to self-interaction. This can indeed be seen from the overly repulsive mean-field electron-electron repulsion potentials predicted by many DFAs [245, 246]. The Li atom is also not multireference, and thus the reason behind the size underestimation by many functionals is not entirely clear. HFLYP not being an outlier with respect to overlocalization indicates that delocalization is unlikely to be a major factor (and bare HF in fact overestimates the size). Nonetheless, most of the modern functionals that fare poorly for H (B97M-V, MN15-L, MN15 and ωB97M-V) continue to fare poorly for Li and overestimate the size, revealing considerable scope for improvement. The size of the Li atom is thus another reasonable choice as a soft constraint for future functional fitting.

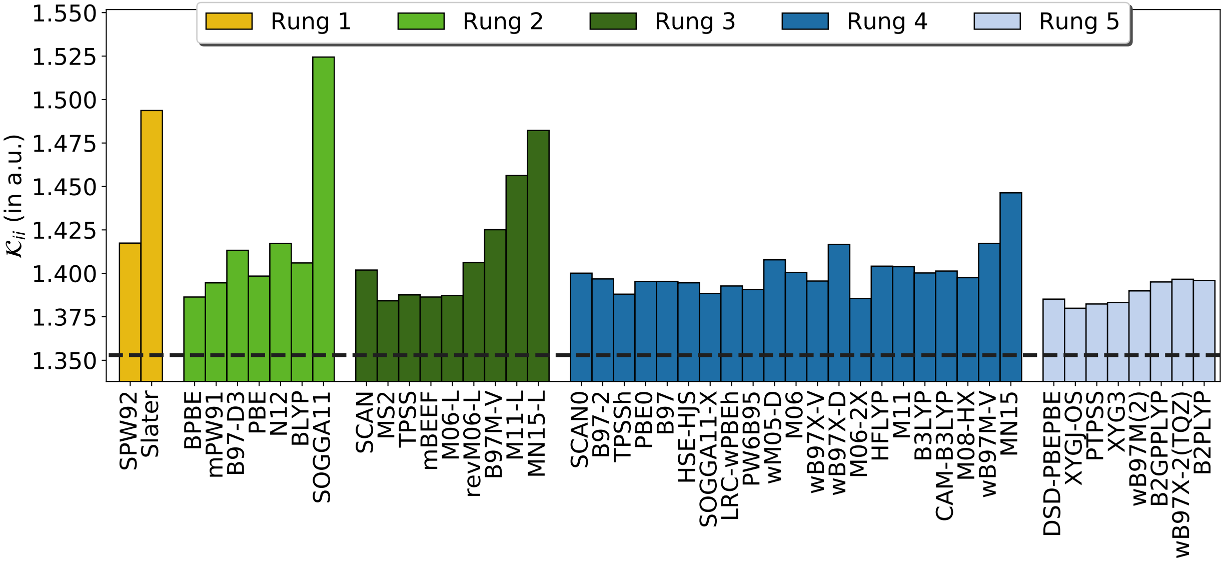
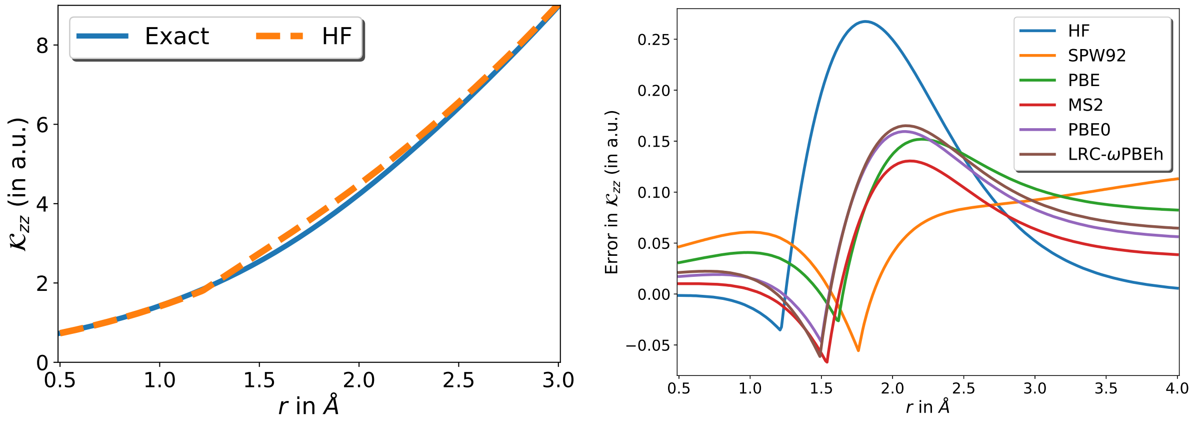


Figure 5.3: $\mathcal{K}_{ii}$ predicted for the Be atom, by various DFAs. The dashed line is the benchmark (CCSD(T)/CBS) value and the HF value is 1.444.

In contrast to the preceding two cases, Fig 5.3 shows that most methods overestimate the size of the Be atom by a similar amount, with some poor performing outliers. In fact, the best mGGA and hybrid have ~ 0.001 variation in $\mathcal{K}_{ii}$ (while the best double hybrid is better by ~ 0.005). This relatively uniform performance is likely a consequence of the challenges faced by DFAs in accounting for the multireference character of the Be atom (it has 0.37 effectively unpaired electrons [247]). It is nonetheless worth noting that the worst performers are recently parameterized modern functionals. It is not entirely clear whether the Be atom should feature in datasets used for functional training because of the partial multireference character, but it should definitely be employed in test sets to gauge fit quality.

5.5 Stretched H_2

The behavior of DFAs in predicting $\boldsymbol{\mathcal{K}}$ in non-equilibrium configurations is potentially interesting. We only investigate the case of H_2, to avoid qualitative failures associated with



(a) Exact $\mathcal{K}_{zz}$ values, along with HF.

(b) Errors in $\mathcal{K}_{zz}$ predicted by various DFAs.

Figure 5.4: Behavior of $\mathcal{K}_{zz}$ for stretched H_2.

delocalization error for polar bonds [78–81] for general bond dissociations. We furthermore restrict ourselves to a few well-behaved functional at Rung 4 and below, to avoid incomplete spin-polarization based problems [228] . Double hybrid functionals are not considered due to N-representability problems in UMP2, [203] and divergence of RMP2 with bond stretching (indicating that they are quite unsuitable for bond dissociation problems). Fig 5.4a shows that $\mathcal{K}_{zz}$ grows quadratically with the internuclear distance r, and Fig 5.4b thus only plots the error in $\mathcal{K}_{zz}$ for clarity. All of the single-reference methods encounter a derivative discontinuity at the Coulson-Fischer (CF) point [212] (which is the onset of spin-polarization) due to the orbital stability matrix being singular at that point. The behavior right beyond this point is the most interesting, as the two electrons are fairly strongly correlated in this regime. The DFAs other than SPW92 yield remarkably similar behavior in this region, systematically overestimating $\mathcal{K}_{zz}$ somewhat (and thus predicting more 'atom' like densities as opposed to a still partially bonded molecule). HF exhibits the same qualitative behavior, but has much larger errors due to complete absence of correlation (indeed, the systematic overestimation past the CF point can be clearly observed in Fig 5.4b as well). Similar overestimation in $\boldsymbol{\mathcal{K}}$ is seen for the Be atom, which also has partial multireference character. Consistent systematic overestimation of $\boldsymbol{\mathcal{K}}$ by several well-behaved functionals with varying levels of delocalization error could in fact potentially be a signature of strong correlation, as the electrons are forced to be less compact in the absence of explicit (i.e. post mean-field) correlation that allows them to occupy the same region of space.

5.6 Discussion and Conclusion

The objective of this work was to characterize the electron density of small molecules using a scalar metric that goes beyond the first moment of the electron density (i.e. dipole moments).

Since the molecular quadrupole moment is origin-dependent for molecules with non-zero charge or dipole, we elected instead to characterize the 3×3 matrix of second cumulants, or spatial variances, of the electron density, $\boldsymbol{\mathcal{K}}$. The eigenvalues of $\boldsymbol{\mathcal{K}}$ are thus a measure of the square of the characteristic extent of the molecular density along each principal axis.

We produced a benchmark dataset, Var213, that contains 213 benchmark values of $\boldsymbol{\mathcal{K}}$, from 100 small molecules, evaluated with CCSD(T)/CBS correlation based on extrapolation towards the complete basis set limit from aug-cc-pCVTZ and aug-cc-pCVQZ calculations, combined with HF/aug-cc-pCV5Z. These reference values were then used to assess $\boldsymbol{\mathcal{K}}$ predictions from 47 density functional approximations (DFAs). Broadly, the molecules studied here are relatively straightforward in terms of their electronic structure. In the language of ref. 55, this dataset should be considered "easy" for DFAs, rather than "difficult" due to the absence of strong correlation or delocalization effects.

The results show that it is possible to obtain quite reasonable $\boldsymbol{\mathcal{K}}$ values in many cases, even with modern density functionals. However, non-empirical functionals, especially at the mGGA level via SCAN, and at the hybrid level via SCAN0, seem to fare better than more empirically parameterized models. This constitutes useful independent validation of the quality of electron densities from these functionals. Double hybrid rung 5 functionals yield the best overall performance: indeed all double hybrids tested match or exceed the performance of SCAN, and only the early B2PLYP double hybrid fails to outperform SCAN0.

Interestingly, the performance of the modern MN15-L, MN15, B97M-V and ωB97M-V functionals over the studied dataset is disappointing, suggesting that the vastness of the meta-GGA space [63] has offered these functionals the opportunity to predict reasonable energies from fairly poor densities. There is an interesting lesson to be drawn from this outcome, particularly for the minimally parametrized meta-GGA-based functionals such as B97M-V and ωB97M-V whose 12-16 parameter form was inferred from the use of large nunbers of chemically relevant energy differences. Specifically, such forms are evidently not fully constrained by the data used to generate them. Independent data such as $\boldsymbol{\mathcal{K}}$ may indeed be very useful in attempting to develop improved data-driven functional designs.

Overall, our $\boldsymbol{\mathcal{K}}$ benchmark tests show that the performance of the best-performing functional strictly improves on ascending Jacob's ladder, indicating that extra complexity has the potential to improve behavior. Use of this dataset and other density based information (and especially the H, Li and Be atoms) for future functional development could thus provide sufficient soft constraints to yield improved $\boldsymbol{\mathcal{K}}$ values, and by implication electron densities as well. The results obtained here may thereby contribute to offering a route to better computationally tractable approximations to the exact functional.

5.7 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: geometries, computed values and analysis.

Chapter 6

Static Polarizabilities

6.1 Introduction

The vast majority of present day electronic structure calculations employ some flavor of Kohn-Sham (KS) density functional theory (DFT) [48, 55, 90, 137]. The exact functional mapping ground state electron density to ground state energy however remains computationally inaccessible, despite proof of its existence [41]. This has led to development of hundreds of density functional approximations (DFAs) of varying complexities through strategies ranging from a sole focus on satisfaction of exact physical constraints [57, 61, 138] to exclusively fitting to benchmark data [120, 175] . Empirically fitted functionals are expected to be extremely accurate within the datasets employed for fitting, but they can be overparametrized to a point where performance outside of that comfort zone becomes rather suspect [80, 115, 131, 211]. Conversely, functionals with none or few fitted parameters may not be as accurate for specific datasets but are believed to be much more transferable to systems not employed in the development process. Many modern functionals have consequently been developed through approaches that draw from both extremes [62, 96], in the hope of having an acceptable balance between fit accuracy and transferability.

In general, a DFA should be transferable to any system of interacting electrons subjected to some external one particle potential $v(r)$. Isolated molecules often employed in fitting functionals satisfy the above condition (with $v(r)$ being the electrostatic potential of the nuclei) but they represent only a small fraction of systems where DFT is applicable. Molecular systems studied with DFT in fact often interact with external fields in chemically relevant scenarios—such as light-matter interactions or in condensed phase. The behavior of density functionals outside the traditional zero external field conditions thus has direct relevance to simulations of spectroscopy, dielectric solvation models [248] and QM/MM, amongst others.

The simplest possible external field is a constant electric field $\vec{\mathcal{E}}$, and the corresponding responses of the energy $E(\vec{\mathcal{E}})$ are well known molecular properties [249]. The first response of the zero-field energy $E(0)$ to $\mathcal{E}$ is the dipole moment $\vec{\mu} = -\left(\dfrac{\partial E}{\partial \vec{\mathcal{E}}}\right)_{\vec{\mathcal{E}}=0}$, which is a simple global measure of the accuracy of electron density in a polar molecule. We studied the accuracy of DFT in predicting $\vec{\mu}$ in an earlier work for the purpose of gauging the quality of DFT densities [80], and will not consider it further here.

The dipole however contains no information about the response of the ground state wavefunction to the electric field as it can be computed from the zero-field density alone. The first property to contain information about the wavefunction response is the static polarizability tensor $\boldsymbol{\alpha}$ which is the second response of the zero field energy to the external field. The components α_{ij} are given by $-\left(\dfrac{\partial^2 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j}\right)_{\vec{\mathcal{E}}=0}$, which allows us to express it in terms of excited electronic states at zero external field using perturbation theory. Specifically, given a ground state $|\Psi_0\rangle$ with energy E_0 and excited states $\{|\Psi_k\rangle\}$ with energies $\{E_k\}$ at $\vec{\mathcal{E}}=0$, we have:

$$\alpha_{ij} = 2\sum_{k=1} \frac{\langle\Psi_k|\,\hat{\mu}_i\,|\Psi_0\rangle\,\langle\Psi_0|\,\hat{\mu}_j\,|\Psi_k\rangle}{E_k - E_0} \tag{6.1}$$

where $\hat{\vec{\mu}}$ is the dipole operator. The ability of a functional to predict static polarizabilities therefore not only reveals how the electron density responds to an external field, but also gives indirect information about the treatment of dipole allowed excited states by the functional. This is a coarser measure than linear-response TDDFT [250] in that the information about a single excited state cannot be disentangled, but it is formally exact as it does not involve use of the adiabatic local density approximation (ALDA) [251]. Polarizability predictions therefore indirectly measure both the transferability of a functional and the quality of excited state predictions (within the limits of the Hohenberg-Kohn [41] theorem), making it useful for assessing functional quality. Fitting functionals to polarizabilities should also assist in ensuring correct response of the energy to external electric fields.

There however does not appear to be a large database of benchmark static polarizabilities which could be used for either fitting density functionals to, or for assessing their performance. Past assessments [147, 252] have often utilized experimental values, which may contain nuclear quantum effects that no *accurate* electronic structure method should reproduce within the Born-Oppenheimer approximation. Some studies [214, 253] also appear to suggest that many experimental polarizabilities may contain substantial errors, making a comparison to benchmark wavefunction theories necessary for a truly fair assessment of the accuracy of polarizability predictions from DFAs. There have been several such studies [254–257], but they appear to either consider only a narrow range of molecules or use an apparently inadequate level of theory for reference values. The closest approximation to a benchmark database of polarizabilities appears to be CCSD(T)/aug-cc-pVTZ values for 145 organic molecules compiled by Wu, Kalugina and Thakkar for assessing the performance of 34 DFAs [258]. We however feel that these values may not be of benchmark quality on account of basis set incompleteness errors, which can be substantial for dipole moments [80] and polarizabilities (as will be shown later). The absence of split core orbitals also leads to an incomplete description of core electrons, and such effects could be significant (as they are already known to matter for dipole moments [148]). Ref 258] also only considered closed shell molecules and did not assess the performance of double hybrid functionals, which are known to be excellent for predicting energetics [111] and densities [80, 259].

We have consequently developed a database of the diagonal components of static polarizabilities of 132 species (composed of main group elements lighter than Ar) at equilibrium geometry, using coupled cluster singles and doubles with perturbative triples (CCSD(T)) [23], extrapolated to the complete basis set limit (CBS). 75 of the species have a stable spin-unpolarized Hartree-Fock (HF) solution while the remaining 57 have a stable spin polarized HF solution either due to open-shell character or spin-symmetry breaking. The not spin-polarized (NSP) subset is expected to be relatively well described by conventional electronic structure methods, but the spin-polarized (SP) subset has the potential to be more challenging. This database is used to assess the performance of 60 popular and recent density functionals, along with HF, second order Møller-Plesset perturbation theory (MP2) and coupled cluster singles and doubles (CCSD). This dataset should also be useful for future functional development and assessment. In particular, it permits evaluation of the accuracy with which various DFAs predict individual diagonal elements of the static polarizability tensor, instead of just their average (where some cancellation of errors may occur) as is the case in 258. The high computational cost of large basis set coupled cluster calculations however constrained us to only consider species smaller than most studied in Ref 258, though we feel general features from this dataset ought to be transferable to larger systems where KS-DFT can be applied. It is however quite likely that somewhat different behavior is observed for systems with substantial multi-reference character, like long chains where the accuracy of KS-DFT polarizability calculations have long been suspect [260, 261].

6.2 Computational Methods

All the calculations were done using a development version of Q-Chem 5 [191]. Most of the equilibrium geometries employed in the study were either obtained from experimental values in the NIST Computational Chemistry Database [192] or were optimized with MP2/cc-pVTZ if experimental geometries were unavailable. The exceptions to this general principle are BH_2F and BH_2Cl, whose experimental geometries were taken from Refs 193 and 194 respectively, as well as the non-covalent complexes FH-OH and H_2O-Li, whose geometries were obtained from the TA13 [197] database. All the geometries employed in the present study are provided in the Supporting Information, along with the source.

Polarizabilities were obtained from finite differences using a central three point formula $\alpha_{ii} = \dfrac{E(\mathcal{E}\hat{i}) + E(-\mathcal{E}\hat{i}) - 2E(0)}{\mathcal{E}^2}$, which should be correct to $O(\mathcal{E}^2)$. The field strength $\mathcal{E}$ was set to 0.01 a.u. for most species, but a few s block atom containing species with large polarizabilities appeared to have large higher-order responses (made evident by a $> 0.5\%$ shift in HF polarizability estimates on changing $\mathcal{E}$ from $0.01 \rightarrow 0.005$ a.u.) that necessitated use of $\mathcal{E} = 0.001$ a.u. for that subset. This smaller field however was not universally employed due to larger risk of contamination from numerical error, especially in species without a highly polarizable s block atom where the HF polarizability estimate was shifting very little on halving the field strength to 0.005 a.u. anyways ($< 0.2\%$ for the vast majority). A full list of field strengths used for each molecule is given in the Supporting Information.

HF polarizabilities were obtained from spin unrestricted calculations with the aug-cc-pCV5Z [70, 198–200] basis, which appeared to be close to the CBS limit (0.2% RMS deviation versus the equivalent aug-cc-pCVQZ numbers). Stability analysis was performed at the aug-cc-pCVQZ level to ensure all solutions were at a minima. Spin unrestricted DFT calculations were done with the aug-pc-4 [65–69] basis for functionals spanning Rungs 1-4 in Jacob's Ladder and it was assumed that such 5ζ basis results ought to be essentially at CBS as well (though behavior for dipole moments suggest that this may not strictly be true for all functionals [80]). Local xc integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points for all atoms, while nonlocal VV10 [107] correlation was calculated over an SG-1 [201] grid (which consists of a subset of points employed in a grid with 50 radial and 194 angular points). Stability analysis was done at the aug-pc-2 level to determine which SCF solutions were potentially unstable, and the problematic aug-pc-4 cases were reoptimized to ensure that the energy was at a minimum.

No orbitals were held frozen for any correlated wavefunction theory (WFT) calculation. All coupled cluster calculations (CCSD/CCSD(T)) were done employing unrestricted orbitals, but known N-representability violations [80, 203] in unrestricted MP2 led us to calculate both unrestricted (henceforth referred to as MP2) and restricted (RMP2) values for comparison. The CBS limit was obtained by extrapolating the correlation component of α_{ii} from aug-cc-pCV5Z and aug-cc-pCVQZ results for the smaller species; and from the aug-cc-pCVQZ and aug-cc-pCVTZ results for the remainder of the dataset. The extrapolation was done via:

$$(\alpha_{ii})_n^{\text{corr}} = (\alpha_{ii})_\infty^{\text{corr}} + A/n^3 \tag{6.2}$$

which is known to be accurate for dipole moments [80, 202]. Eqn 6.2 was found to be accurate to $\approx 0.1\%$ in predicting aug-cc-pCV5Z polarizabilities of some of the smaller species from aug-cc-pCVQZ and aug-cc-pCVTZ results, indicating that it was reasonably accurate for polarizabilities as well.

The superior performance of RMP2 over unrestricted MP2 led us to only perform calculations with restricted orbitals for double hybrid functionals, with no orbitals held frozen here as well. The xc integrals for double hybrid functionals were calculated using the same grids as all other functionals. The non-MP2 component of α_{ii} for double hybrid functionals was assumed to be essentially at the CBS limit when the aug-cc-pCV5Z basis was employed ($0.2 - 0.3\%$ RMS deviation versus the equivalent aug-cc-pCVQZ numbers). The MP2 correlation component was extrapolated to CBS from aug-cc-pCVQZ and aug-cc-pCVTZ results, using Eqn 6.2. Calculations employing the XYGJ-OS [189] functional were accelerated using the RI approximation [204] with the riMP2-cc-pVTZ auxiliary basis [205, 206] for aug-cc-pCVTZ calculations and riMP2-cc-pVQZ auxiliary basis [206] for aug-cc-pCVQZ calculations.

The error in α_{ii} against the reference value $\alpha_{ii,\mathrm{ref}}$ was defined to be $\dfrac{\alpha_{ii}-\alpha_{ii,\mathrm{ref}}}{\alpha_{ii,\mathrm{ref}}}\times 100\%$, which is a purely relative error. This was used to calculate root-mean squared, mean and maximum absolute errors for each of three diagonal components of the $\boldsymbol{\alpha}$ tensor, which are listed in the Supporting Information. The components themselves however have very similar relative error values, and so we only report the root mean square relative error (RMSRE), mean relative error (MRE) and maximum magnitude relative error (MAX) *over all the species and over the three diagonal components*, in the manuscript. In other words, if $\epsilon_{i,l}$ is the error in α_{ii} for species l, we have:

$$\mathrm{RMSRE} = \sqrt{\frac{1}{3L}\sum_{i=x,y,z}\sum_{l=1}^{L}\epsilon_{i,l}^2} \tag{6.3}$$

$$\mathrm{MRE} = \frac{\sum_{i=x,y,z}\sum_{l=1}^{L}\epsilon_{i,l}}{3L} \tag{6.4}$$

$$\mathrm{MAX} = \max\left(|\epsilon_{i,l}|\forall i \in \{x,y,z\}\,\mathrm{and}\,1 \leq l \leq L\right) \tag{6.5}$$

where L is the total number of species. The difference between the RMS relative errors for the α_{ii}'s with the largest and smallest RMS errors (DIFF) is also reported, in order to determine if performance varies considerably along different axes.

6.3 Results

The error metrics defined in the preceding section were used to evaluate the performance of HF, MP2, CCSD and 60 DFAs against the CCSD(T) benchmark. The errors for all the electronic structure methods are given in Table 6.1, while while Fig 6.1 provides a visual representation of performance for a subset of the methods. The full list of all errors for all methods is provided in the Supporting Information, along with all computed α_{ii} polarizability components.

6.3.1 Performance of Wavefunction Theory

CCSD proves to be the most effective method for predicting dataset polarizabilities overall, producing an RMSRE of 1.62% and giving very similar performance over the NSP and SP subsets. This is partly on account of a conscious choice to only include species with maximum deviation less than 7% between CCSD and CCSD(T) α_{ii} in the dataset, in the hope that this would restrict the dataset to species where higher order excitations like exact triples or quadruples would not be necessary to obtain benchmark quality numbers. A large portion of the CCSD error appears to stem from systematic underestimation (MRE is -1.2%) due to the absence of correlation energy from connected triples. It is therefore possible that CCSD

Method	**Class**	**RMSRE**			**MRE**	**MAX**	**DIFF**	**Method**	**Class**	**RMSRE**			**MRE**	**MAX**	**DIFF**
		All	**NSP**	**SP**						**All**	**NSP**	**SP**			
CCSD	WFT	1.62	1.55	1.71	-1.20	6.26	0.34	ωB97X-V [108]	hGGA	4.58	4.02	5.23	2.38	20.12	0.39
MP2	WFT	11.12	2.17	16.68	-0.28	134.01	16.14	SOGGA11-X [162]	hGGA	4.65	3.39	5.88	1.50	27.71	0.42
HF	WFT	8.45	6.84	10.15	-3.04	58.99	3.21	rcamB3LYP [186]	hGGA	4.67	4.10	5.31	1.37	24.07	0.73
RMP2	WFT	8.14	2.17	11.94	0.71	197.48	15.07	B97-2 [118]	hGGA	4.96	3.88	6.08	3.05	25.13	0.08
RHF	WFT	7.60	6.84	8.48	-2.81	30.37	1.56	camB3LYP [94]	hGGA	4.97	4.18	5.85	2.67	20.21	0.11
								LRC- ωPBE [178]	hGGA	5.07	4.84	5.35	3.05	21.63	0.69
SPW92 [138, 149]	LSDA	11.93	10.61	13.31	8.57	119.87	4.22	ωB97X-D [165]	hGGA	5.18	4.13	6.29	3.37	21.44	0.50
Slater [138]	LSDA	19.15	19.39	18.83	16.74	102.68	1.99	B3LYP [89]	hGGA	6.24	5.37	7.20	4.18	29.50	0.39
								B97 [175]	hGGA	6.40	4.91	7.88	4.19	49.07	1.28
N12 [159]	GGA	8.92	7.93	9.98	5.72	48.35	1.60	HFLYP [151]	hGGA	8.61	8.61	8.60	-6.21	34.30	1.36
MPW91 [262]	GGA	9.20	8.70	9.81	7.22	62.06	1.37								
B97-D [59]	GGA	9.48	8.57	10.56	7.40	53.07	0.46	SCAN0 [174]	hmGGA	3.98	2.60	5.23	0.39	18.73	0.76
PBE [57]	GGA	10.11	9.64	10.69	7.91	68.07	1.73	ωM05-D [168]	hmGGA	4.33	3.73	5.00	2.31	19.89	0.09
BLYP [58, 151]	GGA	12.14	10.73	13.78	8.94	75.44	1.06	PW6B95 [180]	hmGGA	4.53	3.75	5.48	2.42	19.31	0.11
SOGGA11 [161]	GGA	77.54	24.91	113.80	14.79	1036.85	45.91	MS2h [172]	hmGGA	4.98	3.91	6.06	2.72	29.23	0.59
								MN15 [126]	hmGGA	5.00	3.50	6.47	2.13	20.54	0.34
mBEEF [182]	mGGA	4.36	3.75	5.05	2.02	18.04	0.50	MVSh [173]	hmGGA	5.08	4.17	5.96	-1.83	21.91	0.12
MVS [173]	mGGA	4.59	3.64	5.54	0.57	22.19	0.46	M06-2X [120]	hmGGA	5.22	3.38	6.93	0.34	35.47	0.35
SCAN [62]	mGGA	5.31	4.79	5.91	3.25	17.93	0.37	M06 [120]	hmGGA	5.24	4.71	5.85	2.99	29.79	0.76
M06-L [101]	mGGA	6.06	3.98	7.99	2.10	45.06	1.88	TPSSh [154]	hmGGA	5.44	5.06	5.88	3.97	23.04	0.24
MS2 [172]	mGGA	6.26	5.27	7.30	4.11	35.89	0.75	M11 [123]	hmGGA	6.13	4.64	7.58	1.71	66.69	3.99
revM06-L [183]	mGGA	6.31	5.32	7.33	0.88	41.83	1.12	M05 [169]	hmGGA	6.46	6.12	6.89	3.35	33.30	1.95
TPSS [61]	mGGA	7.39	6.99	7.89	5.76	34.52	0.55	ωB97M-V [96]	hmGGA	6.83	5.55	8.19	3.91	26.32	0.96
B97M-V [63]	mGGA	7.65	6.73	8.68	5.13	31.55	0.06	BMK [92]	hmGGA	9.49	3.00	13.75	1.53	90.11	6.72
MN15-L [127]	mGGA	7.95	7.18	8.84	6.12	43.16	0.12	MN12-SX [163]	hmGGA	10.39	5.35	14.56	3.12	90.29	4.92
t-HCTH [181]	mGGA	8.27	7.39	9.31	6.05	40.83	1.03	M06-HF [121]	hmGGA	10.69	9.65	11.84	0.04	58.68	1.76
MN12-L [125]	mGGA	15.14	7.05	21.52	3.98	200.38	11.27								
M11-L [124]	mGGA	15.41	11.86	19.01	7.65	126.79	3.12	XYGJ-OS [189]	dhGGA	2.44	1.33	3.37	0.70	15.46	0.45
								XYG3 [188]	dhGGA	2.68	1.53	3.68	0.20	14.98	0.44
MPW1K [84]	hGGA	3.27	2.94	3.67	-0.38	17.36	0.85	DSD-PBEPBE-D3 [209]	dhGGA	2.73	1.76	3.63	1.12	14.30	0.25
PBE50 [263]	hGGA	3.71	3.45	4.03	-1.10	18.25	1.00	B2GPPLYP [190]	dhGGA	3.40	2.16	4.55	1.31	20.23	0.24
MPW1PW91 [262]	hGGA	3.86	3.52	4.27	2.39	14.99	0.21	ωB97X-2 [187]	dhGGA	3.60	2.89	4.37	1.79	16.89	0.46
BHHLYP [58, 151]	hGGA	3.95	2.92	4.99	-0.17	22.08	0.18	B2PLYP [97]	dhGGA	4.36	3.10	5.60	2.49	27.45	0.36
LRC- ωPBEh [179]	hGGA	4.28	3.93	4.68	2.51	18.28	0.63								
PBE0 [91]	hGGA	4.29	3.98	4.67	2.84	16.19	0.28	PTPSS-D3 [208]	dhmGGA	3.59	2.89	4.23	2.05	25.63	0.28
HSEHJS [177]	hGGA	4.31	4.00	4.69	2.86	16.99	0.26	PWPB95-D3 [208]	dhmGGA	3.72	2.54	4.85	1.81	24.37	0.39

Table 6.1: RMS relative errors (RMSRE), mean relative errors (MRE) and maximum magnitude relative errors (MAX) for electronic structure methods over the dataset, expressed as percentages. The RMSRE of the spin-polarized (SP) and non-spin polarized (NSP) subsets of the dataset are also reported separately. Additionally, the difference between the RMS relative errors of the components α_{ii} with the largest and smallest RMS relative errors is reported under DIFF. WFT stands for wavefunction theory, LSDA for local spin density approximation, GGA for generalized gradient approximation, mGGA for meta-GGA, hGGA for hybrid GGA, hmGGA for hybrid meta-GGA, dhGGA for double hybrid GGA and dhmGGA for double hybrid meta-GGA.

(and potentially even CCSD(T)) may not be adequately accurate in predicting benchmark quality polarizabilities outside of our carefully curated dataset, necessitating incorporation of higher order excitations into the coupled cluster scheme. The closeness between CCSD and CCSD(T) values for our dataset, however, indicate that this is not likely to be a concern in the present case.

MP2 performs very well over the NSP species, yielding a small RMSRE of 2.17% that compares very favorably to the equivalent CCSD RMSRE of 1.55%. This behavior is consistent with the excellent performance of MP2 in predicting polarizabilities of closed shell organic molecules noted in Ref 258. N-representability violations [203] in unrestricted MP2 however leads to catastrophically poor performance (including a few negative α_{ii}'s!) over the SP dataset, resulting in a rather high RMSRE of 11.12% over the entire dataset. Switching over the restricted MP2 (RMP2) only partially ameliorates the problem by bringing the RMSRE down to 8.14%, which is still quite large relative to many density functionals. The large SP subset RMSRE likely stems from RMP2 operating on a restricted HF reference that is higher in energy than the spin-polarized solution, which leads to extremely large errors for some species like O_2 (predicted $\alpha_{zz} = 6.732$ Å^3 vs benchmark value of 2.263 Å^3)–despite lack of N-representability violations. MP2 therefore is likely to be a poor choice in general for predicting polarizabilities of spin-polarized species, although it remains an excellent choice for spin unpolarized species.

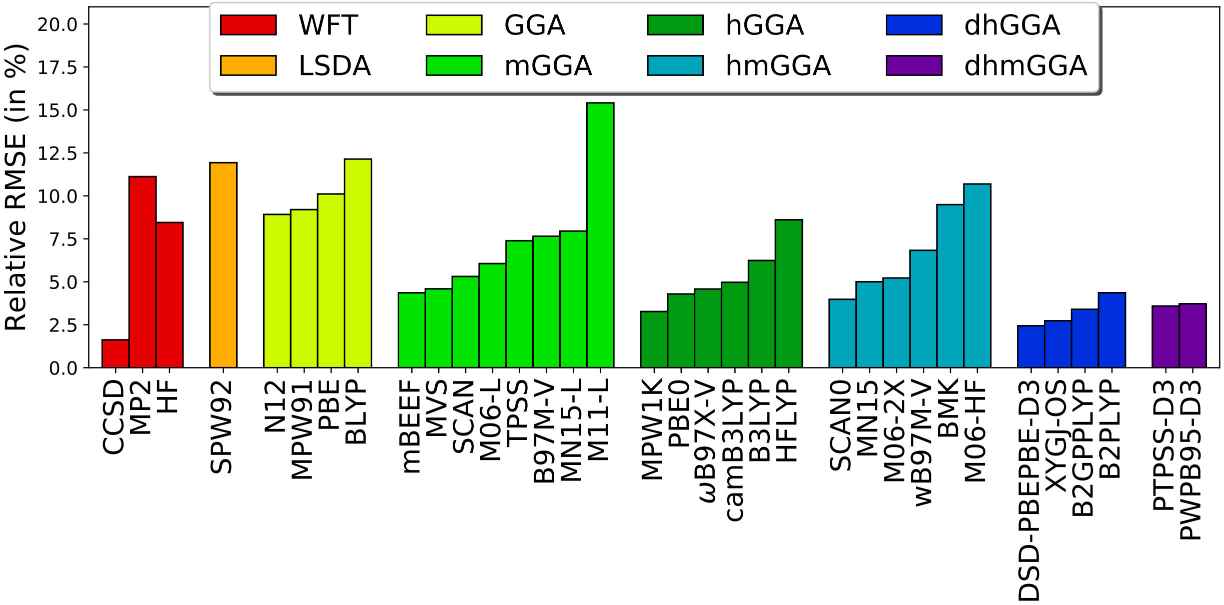


Figure 6.1: Visual representation of the RMSRE for a selection of methods listed in Table 6.1

HF gives quite poor performance, yielding an RMSRE of 8.45%. Lack of correlation energy in HF theory typically leads to larger excitation energies (as can be seen in CIS [251]), which leads to larger denominators in Eqn 6.1 and consequently smaller α_{ii}. This systematic underestimation however is likely only part of the problem, as the MRE is a relatively small -3.04% vs the RMSRE, and the extra error likely stems from the poor quality of the method itself. Interestingly enough, restricted HF (RHF) has a slightly smaller RMSRE of 7.6% while the MRE remains essentially unchanged, suggesting that the constraint of spin restriction eliminated some minor sources of non-systematic error. It is also interesting to note that the RHF RMSRE is slightly lower than the RMP2 RMSRE, indicating that the expensive perturbation theory correction did little to help for SP species.

6.3.2 Performance of DFT

Table 6.1 reveals a number of general features that are likely to be transferable beyond our specific dataset. It is firstly striking that all but five functionals have a positive MRE. This systematic overestimation of polarizabilities for most functionals can be rationalized in terms of self-interaction error resulting in softer, more polarizable densities. The positive MRE can also interpreted in terms of Eqn 6.1: most DFAs tend to underestimate excitation energies (with both TDDFT [251] and to a lesser extent, ΔSCF [264] type approaches [265]), leading to a smaller denominator in Eqn 6.1 and consequently larger polarizabilities. The values in Table 6.1 are consistent with this interpretation, with four of the five functionals producing negative MRE containing 40% or more exact exchange. MVSh [173] is the exception as it contains only 25% exact exchange, but this is likely on the account of the base MVS [173] functional already having a very low MRE of 0.57% despite being a local functional.

It is also important to note that the best DFA in each Rung of Jacob's ladder has a lower RMSRE than the best DFAs in the Rungs below, which is likely a consequence of the additional physical content at each Rung. However, not all functionals employ that extra physical content optimally, as can be seen from the best Rung 3 functionals like mBEEF [182] and MVS outperforming widely used hybrid functionals like B3LYP [89] and M06 [120]. There is in fact substantial overlap in accuracy between Rungs 3 and 4, indicating that exact exchange is not always necessary to improve polarizability predictions.

The best functionals for polarizability predictions are however the double hybrids XYGJ-OS [189] and XYG3 [188] (RMSRE of 2.44% and 2.68% respectively), with DSD-PBEPBE-D3 [209] not far behind (2.73% RMSRE). All three of these functionals were found to be excellent for predicting dipole moments as well [80], indicating that they might be good options for predicting electric field responses in general. They also have a much lower NSP RMSRE ($< 2\%$) than SP RMSRE ($> 3\%$), indicating that use of restricted orbitals likely prevented them from realizing their full potential. This, in conjunction with other failures of conventional double hybrid functionals for spin polarized systems [80, 81] seems to suggest that it is desirable to develop double hybrid functionals based on orbital optimized MP2 type approaches [266, 267] in order to take full advantage of their complexity. It is nonetheless interesting to note that the SP RMSRE for most double hybrids (3-5%) is comparable to, or lower than the SP RMSRE for most hybrid functionals (despite the latter using unrestricted orbitals) and mark a substantial improvement over RMP2. It is therefore likely quite safe to use double hybrid functionals with restricted orbitals for spin-polarized systems at equilibrium geometry, unlike RMP2 (which gives quite poor performance for some species like O_2).

MPW1K [84] proves to be the best Rung 4 functional by far with 3.27% RMSRE–partially as a consequence of the functional employing a large fraction of HF exchange (42.8%), which eliminates the systematic overestimation of polarizabilities present in the related hybrid MPW1PW91 [262] (25% HF exchange) and the MPW91 [262] GGA. In fact, MPW1K has a small bias towards underestimating polarizabilities (MRE of -0.38%). A similar trend is also seen in the case of the second best performing hybrid functional, PBE50 [263] (50% HF exchange), relative to the related hybrid PBE0 [91] (25% HF exchange) and the PBE [57] GGA. The higher HF exchange fractions present in these functionals however could worsen predictions of other properties (as can be seen from the mediocre performance of PBE50 in predicting dipoles, relative to the excellent PBE0 functional [80]). The excellent performance of MPW1K in predicting polarizabilities could also partially stem from the improved accuracy of the base MPW91 GGA relative to PBE and BLYP, resulting in superior predictions by MPW1K relative to other large HF exchange containing functionals like PBE50 and BHHLYP, and could also explain why MPW1PW91 outperforms the classic PBE0 functional. Other hybrid functionals like ωB97X-V [108], ωM05-D [168], PW6B95 [180] and PBE0 also yield quite good performance with around 4.5% or so RMSRE, and are also excellent for dipole moments [80], making them safe choices for estimating electric field responses in general. There are also several others that have errors in the 5-7% range. Hybrid functionals overall also have somewhat larger RMSRE over the SP subset than the NSP one, showing that the former are more challenging. The performance gap however is often quite small ($<$1%), indicating that this is not a major issue for many functionals. BMK [92] however proves to be a major exception, with the SP species increasing the total RMSRE to 9.49% relative to the excellent NSP subset RMSRE of 3%. BHHLYP, SCAN0, B97-2 [118], SOGGA11-X [162] and MN15 [126] are also impacted by the NSP/SP performance difference, though to a much lesser extent. It is also interesting that hybrid meta-GGAs (hmGGAs) do not appear to outperform hybrid GGAs (hGGAs), despite meta-GGAs (mGGAs) being much more accurate than GGAs. This suggests that the vastness of the hmGGA functional space has not been optimally utilized for improving polarizability predictions, which is similar to behavior observed for dipole moments [80].

It is therefore noteworthy that many mGGAs (mBEEF, MVS and SCAN in particular) are essentially as good as hybrid functionals at predicting polarizabilities, which is quite unlike the behavior observed for dipole moments [80]. These three functionals also do not have particularly large systematic biases relative to hybrid functionals (as reflected by their MRE), indicating that this particular manifestation of self-interaction error might be rather minimized. It would therefore be interesting to examine whether these functionals are comparable to hybrids for excitation energy predictions. The case of MVS is rather interesting in particular as its hybrid variant MVSh predicts worse polarizabilities (unlike other similar local/hybrid pairs like SCAN/SCAN0 or TPSS [61]/TPSSh [154]) due to MVS already having a very low systematic bias. MVS was also mediocre in estimating dipole moments, whereas mBEEF and SCAN were the best local functionals tested for dipole predictions [80].

Overall, there is no great chasm in performance between local functionals and hybrids (unlike in dipole moments [80]), with several less accurate mGGAs like MS2 [172] and M06-L [101] giving performance comparable to many hybrid functionals, and some others like TPSS and B97M-V [63] being not much worse. The only potentially problematic aspect is the larger gap between NSP and SP RMSREs relative to hybrid functionals, but the actual errors remain fairly low. There is however a rather wide gulf separating mGGAs from GGAs, with the latter being rather poor at predicting polarizabilities relative to mGGAs. Aside from the special case of SOGGA11 [161] however, GGAs represent an improvement over LSDA but nonetheless ought not be employed to estimate polarizabilities as mGGAs make significantly better predictions for the same computational complexity.

Coming to the matter of poor performers, we find that the worst double hybrid functional B2PLYP [97] actually does not have particularly bad RMSRE (4.36%) relative to Rungs 3 and 4. However, its performance is quite poor relative to its Rung 5 brethren (as was also observed to be the case for dipole moment predictions [80]), and it has a higher RMSRE than the related BHHLYP hybrid functional, indicating that the excess functional complexity did not lead to any evident advantages. Interestingly enough, the PWPB95-D3 functional appears to be one of the worse performing double hybrids, despite being the most accurate of the Rung 5 functionals tested for dipole moment predictions [80]. The worst hybrid functionals give much less satisfactory performance with RMSRE $\approx 10\%$ but are still better than LSDA. The worst GGA (SOGGA11) and mGGA (M11-L [124]) however are worse than Rung 1 LSDA, suggesting overparametrization has led to a subsequent decrease in accuracy. SOGGA11's performance in particular is exceptionally poor for even simple species like H_2 and HCl, to say nothing of more challenging cases like Li_2. This functional also has a rather poor basis set convergence for dipole moments [80], suggesting that the aug-pc-4 numbers may not be adequately close to the CBS limit. SOGGA11 therefore should not be used for molecular property predictions. Paradoxically however, the related SOGGA11-X hybrid functional performs quite well, suggesting that there may exist some scope for improving SOGGA type functionals.

It is also interesting to note that the best performing functionals have small DIFF values ($\approx 1\%$ or lower). MRE for individual α_{ii} are also quite similar in value, suggesting that these functionals are about equally as effective in estimating $\alpha_{xx,yy,zz}$. This is encouraging, as it suggests that good functionals would predict accurate static polarizabilities on account of getting individual components correct, instead of relying on cancellation of errors between them. Several of the poor performing functionals however have quite large DIFF values, largely as a consequence of catastrophic failure in predicting α_{zz} (polarizability along the bond axis) for certain s-block diatomics like Li_2 or NaLi, while getting more reasonable estimates for the components perpendicular to the bond axis.

6.3.3 Challenging Species

The nine most challenging species for DFT polarizability predictions are listed in Table 6.2, roughly in descending order of difficulty. Most are SP species, which is entirely unsurprising on account of the best performing functionals being double hybrids that are constrained to act on restricted references alone. The two most challenging species in the dataset by far are the H and Be atoms, which is somewhat surprising on account of their simplicity. The polarizabilities for both are systematically overestimated by $\approx 10\%$ or more, although

Species	**Subset**
Be	SP
H	NSP
CN	SP
C_2H	SP
H_2O-Li	SP
Li_2	SP
Na	SP
NaLi	SP
NaH	NSP

Table 6.2: The 9 most difficult species in the dataset. These were selected on the basis of the first quartile of relative RMS DFA errors for each species (the RMS being taken over the three components). The selected species give $\geq 5\%$ first quartile error, which corresponded to a break in the distribution. Coincidentally, this number is also comparable to typical Rung 4 RMSREs.

the errors for H decrease on increasing the fraction of exact exchange over families of similar functionals (expectedly, as HF is exact for the H atom), suggesting that self-interaction error is to blame in this case. On the other hand, Be is likely challenging due to multi-reference character (HF breaks spin-symmetry despite Be formally being a closed-shell singlet atom) and not self-interaction, as the errors remain similar over related functionals with different exact exchange fractions. The behavior for both atoms is quite troubling as it indicates that thirty years of functional development involving many comparisons and fits to such simple atoms nonetheless fails to predict response under a constant electric field. We recommend that the zero external field atomic energies typically used for fitting functionals be augmented with energies and densities for atoms subjected to diverse field strengths, in order to generate a more complete description of such so called ‘perfect norms’ [115].

Other challenging species are the triple bonded C_2H and CN radicals, and highly polarizable alkali atom containing species like NaLi. Their presence on the list is unsurprising, especially since the latter poses a major challenge to functionals with regards to density predictions [80]. MN12-L [125] and MN12-SX [163] in particular fail catastrophically in predicting properties for such alkali atom containing species, indicating a need for caution in applying these Minnesota functionals to s block elements. In general however, the species in Table 6.2 are challenging for the vast majority of functionals and no class of DFAs appear to have an evident advantage for any of them. The list of best and worst performing functionals for each species in Table 6.2 contain a mix of both local functionals and hybrids, irrespective of the extent of single/multi-reference character.

It is also interesting to note that the most challenging species in Table 6.2 are not amongst the largest species in the dataset. KS-DFT is known to have significant difficulties in predicting $\boldsymbol{\alpha}$ for long chains [260, 261, 268], although none of the molecules studied were anywhere close to the problematic length scales. The possibility of different behavior on very large length scales therefore cannot be ruled out, especially if the systems have substantial multi-reference character like long polyene or acene chains. For the 132 systems studied in this work however, atoms as a rule prove to be the most challenging for most methods, followed by linear molecules and then non-linear species. The precise rankings of functionals shift only slightly depending on whether atoms, linear molecules or non-linear systems are being studied. A few functionals like BMK and M06-2X exhibit somewhat larger variation in performance on account of catastrophic failure for a few systems (diatomics for the former and atoms for the latter). The overall relative performance of functionals (especially within a given rung) however appear to be independent of the size/geometry of systems. The best functionals consistently yield excellent results for both atoms and molecules, while the worst performers continue to generate poor results.

Method	**RMSRE**		**MRE**		**MAX**	
	CVQZ	**CVTZ**	**CVQZ**	**CVTZ**	**CVQZ**	**CVTZ**
CCSD(T)	0.46	1.33	0.16	-0.12	2.54	9.14
CCSD	0.44	1.19	0.21	-0.02	2.08	8.05
RMP2	0.43	1.42	0.05	-0.37	2.76	9.55
HF	0.19	1.04	-0.03	-0.60	1.43	6.86

Table 6.3: Percentage differences between the CBS estimate and aug-cc-pCV(Q/T)Z α_{ii} for WFT methods. RHF has essentially same basis set convergence as HF, while MP2 basis-set convergence rate is unreasonably slow due to N-representability violations. A full table with all deviations is supplied in the Supporting Information.

6.3.4 Basis Set Convergence

HF polarizabilities converge somewhat slowly (relative to behavior seen for dipole moments) on increasing basis set size. The RMS deviation between aug-cc-pCV5Z and aug-cc-pCVQZ numbers is 0.2%, while the equivalent deviation between aug-cc-pCV5Z and aug-cc-pCVTZ is 1 %. This indicates that the aug-cc-pCV5Z are likely essentially at the CBS limit, but the aug-cc-pCVTZ numbers are not an optimal replacement. No detailed investigation of basis set convergence of DFT polarizabilities were carried out, but it is not unreasonable to expect behavior similar to HF for all but a few ill behaved functionals that are notorious for slow basis set convergence [80, 211]. This would indicate that quadruple zeta basis numbers might be sufficiently close to CBS for practical purposes, but triple zeta basis numbers are likely insufficiently converged.

Correlated WFT estimates for polarizability converge more slowly with increasing basis set size, but does so in a relatively well described manner given by Eqn 6.2. A comparison between the CBS estimate and aug-cc-pCVTZ/aug-cc-pCVQZ numbers (as given in Table 6.3) reveal that neither finite basis is an acceptable estimate, and extrapolation almost essential to have benchmark quality numbers. This further indicates that the triple zeta CCSD(T) estimates in Ref 258 would benefit from some refinement (which ought to also incorporate a more complete description of excitations out of the core).

6.4 Conclusion

In summary, we find that it is possible to get quite accurate static polarizability estimates for small and medium sized species from several DFAs, although most seem to systematically overestimate the values on account of self-interaction error. The best performing functionals are double hybrids, with RMSRE in the range of 2.5-3.8%–placing them within striking distance of CCSD (which has 1.61% RMSRE). A few (XYGJ-OS and XYG3 in particular) are essentially as good as CCSD for non spin-polarized (NSP) species, but their overall performance is worsened on account of being constrained to a restricted reference for even spin-polarized (SP) systems, in order avoid N-representability breakdowns. Double hybrid functionals employing an orbital optimized MP2 like formalism [269] therefore could lead to significant further improvement in prediction of polarizability (and other molecular properties), as such N-representability violations would not be a concern.

Hybrid functionals also give quite decent performance, with several yielding RMSREs between 4-5%. Truly exceptional performance however is obtained from the mBEEF and MVS mGGAs, which give hybrid functional level performance at a much lower computational cost. This suggests that these functionals are able to minimize some aspect of self-interaction error, and indicate they may yield better excited state predictions than what might typically be expected of local functionals. Other mGGAs like SCAN and M06-L also give satisfactory performance, indicating that there is no large difference in performance between Rungs 3 and 4 of Jacob's ladder. It would be interesting to see if similar features transfer over to long carbon chains where large fractions of exact exchange have long been held to be essential for decent polarizability predictions [260, 268] but that is beyond the scope of our present study due to difficulty in getting CBS quality benchmark reference data. There is however a fairly large gap between Rungs 2 and 3, on account of GGAs performing considerably worse than mGGAs. We therefore recommend that only functionals from Rung 3 or higher be used for calculations where polarizabilities matter.

Several of the best performing functionals within each rung (N12 [159], mBEEF, SCAN0, XYGJ-OS etc.) are quite recent, indicating that some measure of progress has been made in improving polarizability predictions over time (which also indirectly suggests an improvement in functional transferability and treatment of excited states). Several of the worst performers like SOGGA11, MN11-L and MN12-SX however are also recent, indicating an inconsistent improvement over time. This is in no small part due to catastrophic failures for a few species, which indicates a lack of transferability of the functional, perhaps as a consequence of overfitting.

It is also quite interesting that the large performance gap between GGAs and mGGAs does not carry over to hGGAs and hmGGAs, which give essentially similar performance.

It is possible that exact-exchange and kinetic energy density correct similar aspects in polarizability predictions, and their combination can do little better. However the hmGGA functional space [63] is vast, and the emphasis till date has been to improve zero-field energetics alone. This has led to functionals like ωB97M-V [96], that are highly accurate for zero-field energetics [55], but somewhat lacking with regards to prediction of properties like dipole moments [80] and polarizabilities. It may therefore be possible to employ information about molecular properties to assist exploration of a larger section of functional space in order to develop hmGGAs that are simultaneously accurate for zero-field energies and properties.

We therefore hope that our dataset of CCSD(T)/CBS benchmark static polarizabilities will assist in development and/or testing of future functionals that predict better energy and density responses to external electric fields, and also provide a better description of dipole allowed excited states. This would ideally assist in developing more transferable functionals that would be better approximations to the exact one. Such functionals would also likely prove useful for studies that go beyond isolated systems in gas phase–such as in simulations of spectroscopy or condensed phase processes. A comparison between the supplied DFT polarizibilities and ones calculated from TDDFT could also prove useful in examining the validity of the adiabatic approximation employed in the latter, which ought to prove useful in improving DFT excited state predictions in general.

6.5 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: geometries, benchmark polarizabilities and table of errors.

Chapter 7

Polarizabilities beyond the Coulson-Fischer Point

7.1 Introduction

Kohn-Sham density functional theory (KS-DFT) [48] is the most widely used electronic structure method used to calculate energies and properties of molecules and extended materials [55, 90, 137]. KS-DFT *assumes* that the exact electron density of a given system can be reproduced by a single Slater determinant (or, in the case of double hybrid functionals, a single Slater determinant with small corrections) and is consequently quite successful for many systems where one Slater determinant heavily dominates the wave function (i.e., single-reference systems). This is not to say that KS-DFT is perfect for all single-reference systems—indeed, delocalization error [64, 74, 75] leads to catastrophic failures in many cases [77, 85, 86, 251, 270]. Nonetheless, these errors are well characterized and can often be mitigated by careful choice of functionals [95].

The performance of KS-DFT is considerably more questionable for multi-reference (MR) systems where many Slater determinants make substantial contributions to the exact wave function, due to large differences between the true kinetic energy and the non-interacting approximation that KS-DFT obtains from a single determinant. However, this also does not imply that KS-DFT cannot be applied to any MR system, especially when unrestricted determinants with broken spin-symmetry are employed. The case of stretched single bonds offers the simplest example: bonded electrons would try to localize on the individual fragments in the highly stretched limit, and a spin-restricted determinant would qualitatively fail to reproduce the density as it would force up and down spins into the same spatial orbitals. However, an unrestricted formalism would permit up and down spins to localize in different spatial locations, allowing a physically correct (albeit spin-polarized) description of the density to arise. Unrestricted single determinant methods consequently break spin symmetry when bonds are stretched beyond a limit called the Coulson-Fischer (CF) point [212]. The determinant however ceases to be an eigenstate of $\hat{S}^2$, as the spin polarized solution would be made out of approximately equal amounts of singlet and triplet (plus smaller intrusions from higher spin states, where possible). This appears to be problematic on the surface, since the true eigenstate for the non-relativistic Hamiltonian (that KS-DFT is attempting to solve for) should be spin pure. However, the spins on the two fragments will be fully noninteracting in the broken bond limit, and the singlet and triplet states of the noninteracting fragments should be equivalent, making this spin polarized state adequate for prediction of relative energies and properties not directly related to spins. Indeed, this is the protocol successfully used to evaluate atomization energies, since it is size-consistent (in the sense that the energy E_{AB} of a system with a highly stretched A—B single bond will asymptote to the sum of energies E_A and E_B of isolated individual fragments A and B respectively). We must however note that efforts have been made to develop 'strong correlation' functionals that operate on spin-restricted densities for treating MR problems without spin contamination issues [110, 271–273]. Unfortunately, such methods are typically not size-consistent [110, 271–273] although they do succeed in reducing the magnitude of the size-consistency energy error substantially relative to traditional spin-restricted methods.

The intermediate regime where the spins on the fragments are not quite independent has more potential to be problematic for spin unrestricted methods. In practice, single reference unrestricted Hartree-Fock (UHF) theory gives qualitatively acceptable results in this regime [14], showing that single reference theories can generate reasonable values even for such fairly MR problems, by taking advantage of the weak interaction between the polarized spins. Unrestricted KS-DFT (UKS) methods, on the other hand, are known to make much more problematic predictions for stretched single bonds, but generally on account of well understood delocalization effects. It is for instance quite well known that the dissociation curves for X_2^+ (where X is any monovalent group like H or CH_3) species will have a barrier for fragment association, and in extreme cases have a *negative* dissociation energy [76] due to overstabilization of fractional charges in the dissociation limit. Similarly, local functionals cause polar bonds to dissociate into fractionally charged constituents [78, 79] and even hybrid functionals delocalize charge too much, resulting in a dipole moment that decays too slowly as the bond is stretched [80].

It is however typically assumed that such delocalization driven unphysical behavior is absent from dissociation curves of nonpolar single bonds in neutral molecules. This is a quite reasonable line of thought as there is no intrinsic driving force towards fractional charges (unlike the case of charged species or polar bonds). It is indeed possible that the bond would dissociate to fragments with fractional spins [109] instead of fractional charges, but this is also rather unlikely as single reference HF and KS-DFT methods tend to overestimate the energy of species with fractional spins [109], biasing the calculation against any such possibility in the dissociation limit. Both UHF and UKS theory thus are commonly expected to dissociate nonpolar single bonds in neutral molecules in a qualitatively correct manner and generate the right dissociation limit of neutral fragments with half-integer spin. This permits widespread use of KS-DFT in studying reactive trajectories, as a reasonable description of bond formation and dissociation processes is expected.

A qualitatively correct description of the stretched bond regime should also ensure reasonable prediction of molecular properties. Indeed, properties like static polarizabilities and force constants should be useful for revealing complications in the description of electronic structure. These properties are second derivatives of the molecular energy with respect to parameters like external electric fields or bond length and should consequently magnify any issues present in the underlying energy/density predictions, that would not be apparent by just looking at the potential energy surface. Such second derivative properties are therefore useful probes of high sensitivity for assessing the quality of electronic structure methods for stretched nonpolar bonds, similar to how the first derivative property dipole moment helps reveal catastrophic density failures for stretched polar bonds [80, 203].

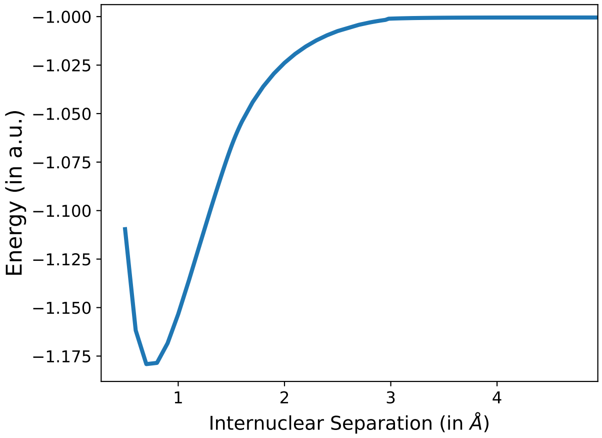


Figure 7.1: Energy prediction for stretched H_2 by TPSS.

Our study of second derivative properties for stretched H_2 reveals that while HF (or more specifically, UHF) and several density functionals like PBE [57] yield reasonable behavior for stretched H_2, many widely used approximations yield unphysical static polarizabilities and force constants, with several also predicting an unphysical barrier for association of two H atoms to form H_2. The underlying potential energy surfaces look quite smooth overall and do not immediately reveal any significant unphysicalities aside from potentially a barrier. Even such barriers are typically < 2.5 % of the bond dissociation energy, and therefore quite hard to detect from a first glance. TPSS [61] for instance has long been claimed to give reasonable energy predictions for stretched H_2 [274], and Fig 7.1 shows it to be quite free of any barriers. The static polarizability and force constants however reveal significant complications between the CF point and the dissociation limit (as is discussed later), indicating that exploration of second derivative properties was essential to identify the points of failure.

In the next section, we detail the computational methods employed by us to reach our conclusions. We subsequently discuss what we believe constitutes reasonable behavior for a KS-DFT method for stretched single bonds, and show that problematic features in static polarizability and force constants are predicted by both empirically fitted functionals like B97-D [59] and M06-L [101], as well as nonempirical ones like TPSS [61], indicating that overparameterization alone is only one potential origin of the issue. We then highlight that these second derivative property prediction failures are connected to unphysicalities in the smallest eigenvalue of the Hessian of the energy with respect to orbital rotations, and incomplete spin localization beyond the CF point. The latter suggests that a fractional bond is preserved even at large internuclear separation, adding a substantial 'ionic' contribution to the UKS wave function, instead of the expected 'covalent' behavior. Afterwards, we show that the use of 'physically correct' (i.e. fully spin polarized) densities could mitigate this problem, and conclude by discussing the barriers towards H atom association for H_2 formation that many functionals predict, suggesting that the problematic functionals ought not to be used to study the *ab initio* dynamics involving bond formation or rupture, especially when better behaved alternatives are available.

7.2 Computational Methods

All calculations were done with the Q-Chem [191] software package. HF/DFT calculations used spin-unrestricted orbitals, the quintuple zeta aug-pc-4 basis [65–67], and an integration grid using 250 radial and 974 Lebedev angular points for the local exchange-correlation integrals, unless specified otherwise. We note that the use of smaller grids or basis sets yielded essentially the same behavior, indicating that the unphysicalities are unlikely to stem from basis set or grid incompleteness errors. Stability analysis was performed for every solution to ensure that the Slater determinant was a (local) minimum with respect to occupied-virtual rotations. The CISD calculations were performed with the aug-cc-pVTZ [198] basis in order to get a qualitative understanding of exact behavior. Static polarizabilities were calculated from a central finite difference formula using an applied electric field strength of 10^{-3} a.u. Exploratory tests with different field strengths yielded the same behavior, indicating negligible finite difference step size errors.

7.3 Results and Discussions

7.3.1 Qualitatively Correct Functionals

We consider a functional to be qualitatively correct if it yields the same general behavior as exact quantum mechanics at all points in the dissociation curve, except possibly at the CF point. Single reference methods can predict non-analytic behavior at the CF point in general [14, 80, 203], and therefore we take HF as the baseline for what constitutes 'reasonable' in the immediate neighborhood of the CF point. This choice stems from HF giving qualitatively right behavior at all other points, and because it seems reasonable to expect UKS to at least fare no worse than the best possible single determinant wave function method (i.e. HF). Indeed, more advanced wave function methods like orbital optimized second order Møller-Plesset perturbation theory (OO-MP2) [266, 267] have to be explicitly regularized to recover HF-like behavior at the CF point [275, 276]. An UKS method that yields behavior closer to the exact results than UHF around the CF point should also be

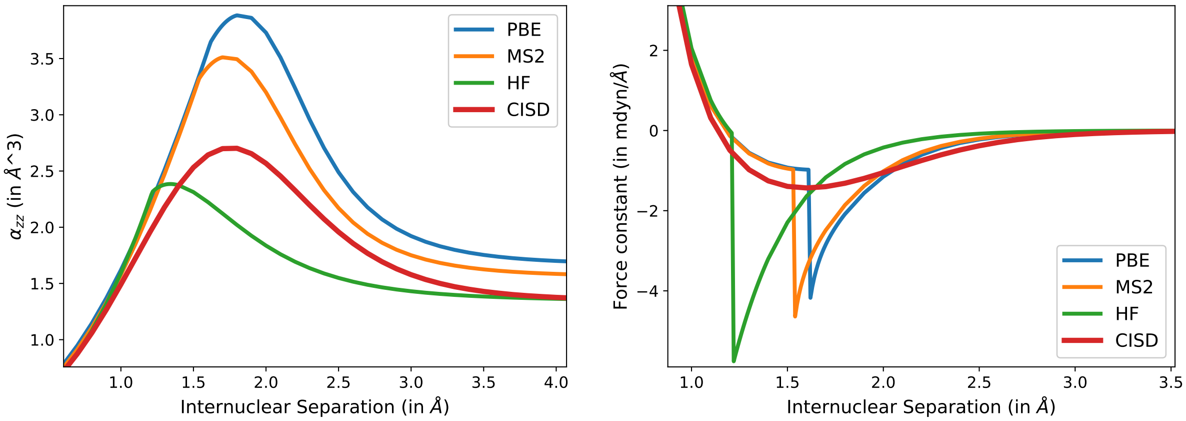


Figure 7.2: Static polarizability along bond axis (left) and force constants for bond stretching (right) predicted by three 'qualitatively acceptable' methods (HF, PBE, and MS2), compared to exact (CISD) values for stretched H_2. The polarizability predictions by all three single determinant methods are qualitatively fine, save a barely perceptible kink at the CF point. There is a discontinuity in force constant predictions at the CF point for all three but the behavior is reasonable otherwise.

viewed as qualitatively accurate in that neighborhood (although failure elsewhere would reflect poorly on the method in general).

Our focus on static polarizabilities and force constants indicates a need to understand the mathematical behavior of properties that can be expressed as second derivatives of the energy. Specifically, given a wave function with a set of orbital degrees of freedom $\boldsymbol{\theta}$, we can say that the derivative of an observable A against some parameter x can be expressed as:

$$\frac{dA}{dx} = \frac{\partial A}{\partial x} + \left(\frac{\partial A}{\partial \boldsymbol{\theta}}\right)_x \frac{\partial \boldsymbol{\theta}}{\partial x} \tag{7.1}$$

from the chain rule, where $\dfrac{\partial \boldsymbol{\theta}}{\partial x}$ represents the rate of change of orbitals against the parameter x. For a variationally optimized energy E, we therefore have:

$$\frac{dE}{dx} = \frac{\partial E}{\partial x} \tag{7.2}$$

as the response of the energy to orbital rotation $\left(\dfrac{\partial E}{\partial \boldsymbol{\theta}}\right)_x = 0$ from the Hellman-Feynman theorem. However, this does not apply to energy second derivatives, which are expressed as:

$$\frac{d^2E}{dxdy} = \frac{\partial^2 E}{\partial x \partial y} + \frac{\partial^2 E}{\partial x \partial \boldsymbol{\theta}} \frac{\partial \boldsymbol{\theta}}{\partial y} = \frac{\partial^2 E}{\partial x \partial y} + \frac{\partial^2 E}{\partial x \partial \boldsymbol{\theta}} \left[\frac{\partial^2 E}{\partial \boldsymbol{\theta}^2}\right]^{-1} \frac{\partial^2 E}{\partial y \partial \boldsymbol{\theta}} \tag{7.3}$$

where $\frac{\partial^2 E}{\partial \boldsymbol{\theta}^2}$ is the Hessian of the energy with respect to orbital rotations (discussed greater detail in Sec 7.5 for the interested reader). This Hessian however has a zero eigenvalue at the CF point of typical single determinant methods [14], due to barrierless degeneracy of the spin-unpolarized restricted and spin polarized solutions. Consequently, second derivative properties may be undefined at the CF point itself due to the inversion of a singular matrix, and is discontinuous in the neighborhood due to discontinuity in $\frac{\partial \boldsymbol{\theta}}{\partial x}$ (the rate of change of orbital degrees of freedom with respect to the parameter x) on both sides of the CF point. This is indeed observed for the UHF force constant, as shown on the right panel of Fig 7.2.

However, the left panel of Fig 7.2 shows that the polarizability (where the applied electric field $\mathcal{E} = x = y$) does not appear to show such a discontinuity at the CF point for H_2, even though a derivative discontinuity (kink) is present. This is a consequence of the lack of a permanent dipole for the molecule. The spin polarization transition therefore does not affect the polarity, ensuring that the rate of change of the dipole $\frac{\partial \mu}{\partial \boldsymbol{\theta}} = \frac{\partial^2 E}{\partial \boldsymbol{\theta} \partial \mathcal{E}} = 0$ along the eigenvector of the Hessian leading to the transition. This eigenvector however is associated with the zero eigenvalue, and therefore no contributions from the singular term survive in Eqn 7.3, leaving behind a continuous function. Polar bonds can have nonzero $\frac{\partial^2 E}{\partial \boldsymbol{\theta} \partial \mathcal{E}}$ along the spin-polarization transition eigenvector and so a discontinuous polarizability can be observed for such systems (and indeed, we observed it for hydrogen fluoride with UHF).

In summary, a derivative discontinuity (kink) at the CF point is acceptable for polarizabilities of nonpolar H_2, while a discontinuity at the CF point is permissible for the force constant. We find that the PBE GGA functional and the MS2 [172] meta-GGA (mGGA) functional satisfy the above criteria for the properties we examine (as can be seen from Fig 7.2), and are therefore taken as a baseline for acceptable behavior for functionals from the second and third rungs of Jacob's ladder [99] respectively. We do however note that these two functionals systematically overestimate polarizabilities on account of delocalization error [95, 226] while HF tends to underestimate polarizabilities in the stretched bond regime due to a systematic bias towards electron localization stemming from missing correlation energy. The hybrid variants PBE0 [91] and MS2h [172] also yield sensible results, as does the popular B3LYP [89] and the related BLYP [58, 151] GGA.

7.3.2 Static Polarizabilities

The willingness of the electron density in the bonded region to 'polarize' is a function of the strength of the chemical bond. A strong bond would be unwilling to distort the natural density and thus have low polarizability. Electrons in a stretched bond are much more likely to be polarizable, as they are weakly held by the nuclei and can be displaced in whichever direction an electric field nudges them to. Beyond a certain point, however, the bonded pair localizes into individual fragments and the polarizability decays to the fragment limit. This behavior is seen for instance in Fig 7.2 (left panel), where the polarizability of H_2 rises with the weakening of the bond via stretching, until the electrons start to localize on individual H atoms, leading to an asymptotic decline to the atomic limit. Reasonable functionals like PBE are expected to give similar results as well, with a kink at the CF point heralding the start of density localization. Unfortunately, we discover that many functionals that are widely used in quantum chemistry yield an inaccurate description of static polarizability for stretched bonds, signaling inaccurate description of the bonding process.

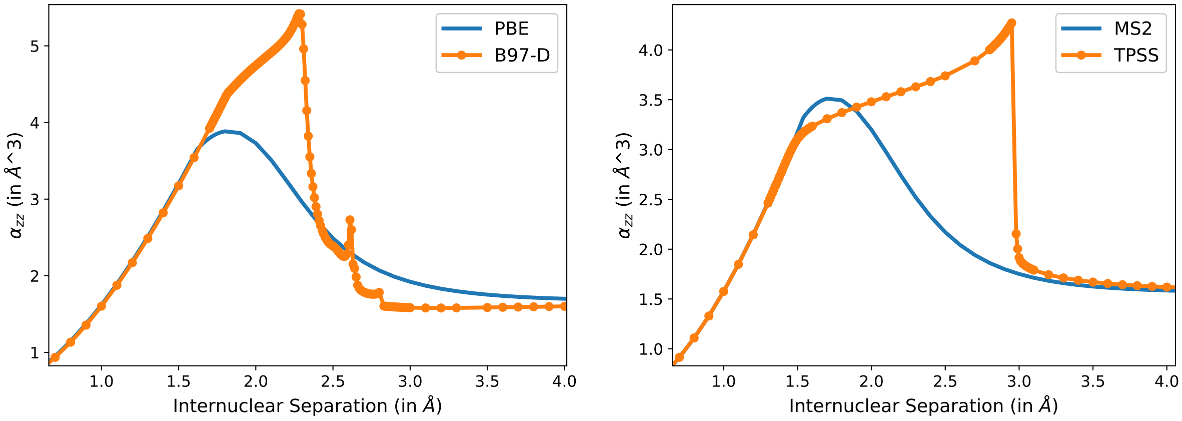


Figure 7.3: Static polarizability along bond axis predicted by B97-D (left) and TPSS (right), with an acceptable functional from the same rung of Jacob's ladder as a reference. Markers for points plotted are given to highlight that the problematic regions were heavily sampled, and so the features are not interpolation/plotting artifacts. The unphysical features here are not local to the CF point region (1.8 Å for B97-D and 1.4 Å for TPSS), although their onset roughly corresponds to that neighborhood.

The behavior of the B97-D GGA and the TPSS mGGA are especially troubling, as they highlight two highly similar and yet distinct modes of catastrophic failure. B97-D is an empirically parameterized functional with 9 parameters, which is regarded as one of the best GGAs for prediction of energetics [55, 111] and is superior to PBE for estimating dipole moments [80] and static polarizabilities at equilibrium geometries [226] . Nonetheless, it predicts monotonically increasing static polarizability well beyond the CF point at 1.8 Å (as can be seen on the left panel of Fig 7.3), indicating that the bonded density was not cleanly localizing to separate fragments. This is followed by a very rapid (though not discontinuous) drop to near the atomic asymptote around 2.3 Å, a second, less prominent peak near 2.6 Å and a small discontinuity at 2.8 Å, suggesting that the UKS density was not being localized into the separate H atoms in a clean, UHF like manner.

TPSS has a similar monotonic increase well beyond the CF point at 1.4 Å, which lasts till about $r = 3$ Å, beyond which it drops discontinuously to a value very close to the atomic asymptote, and afterwards decays smoothly. This transition appears to be discontinuous (as opposed to merely very rapid, which was the case for B97-D) based on sampling at intervals of 0.01 Å, which is quite troubling as it suggests a dramatic, discontinuous change in the underlying density. At any rate, TPSS seems poorly suited for finding electric field responses for H_2 at stretched geometries.

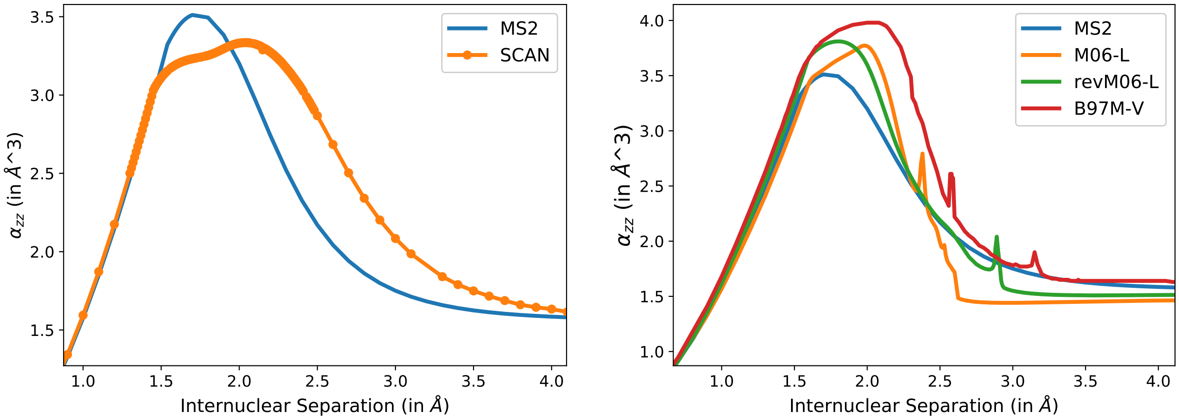


Figure 7.4: Static polarizability along bond axis predicted by SCAN (left) and empirically trained functionals M06-L, revM06-L and B97M-V (right), alongside reference MS2 values.

The more recently developed SCAN [62] functional gives better performance than TPSS, but is not entirely free from faults. The left panel of Fig 7.4 shows that SCAN predicts two very closely spaced peaks in the polarizability curve, which is not quite what one ought to expect. This twin peak structure essentially leads to a plateau instead of a peak in the polarizability values along the dissociation curve, suggesting that the density localization is not happening as quickly as in HF, although is likely still happening in a smooth manner, unlike the case of TPSS. This strange behavior is nonetheless worthwhile to keep in mind while developing future meta-GGA functionals, as it shows that all of the strong nonempirical constraints and appropriate norms employed to develop SCAN were not by themselves sufficient to reproduce reasonable behavior for stretched H_2, which simpler functionals like PBE can readily achieve. It is also important to note that these unphysical features of TPSS and SCAN persist in their hybrid variants TPSSh [154] and SCAN0 [174] (although the polarizability values are lower in those cases), indicating that this behavior is not solely a consequence of some form of delocalization error.

It is also interesting to consider the performance of empirical mGGAs like M06-L [101], revM06-L [183] and B97M-V [63] against SCAN and TPSS. These functionals give reasonable behavior for most parts of the dissociation curve (as can be seen in the right panel of Fig 7.4, and in more detailed figures in the supplemental information), but regions of local failure are very much visible in the form of unexpected local maxima. M06-L also has a small discontinuity around 2.6 Å, while B97M-V appears to have two: one a sudden drop around 2.3 Å and the other the peak around 2.6 Å; suggesting discontinuous changes in density. These features can be seen more clearly in Figures S4 and S8 in the supplemental information.

7.3.3 Force Constants

The failures of these functionals in predicting static polarizabilities at stretched geometries naturally lead to the question as to whether any other second-order property (i.e. based on a second derivative of the energy) shows similar unphysical behavior. The force constant (i.e. the negative of the second derivative of energy with respect to bond stretching) is therefore a natural observable to investigate.

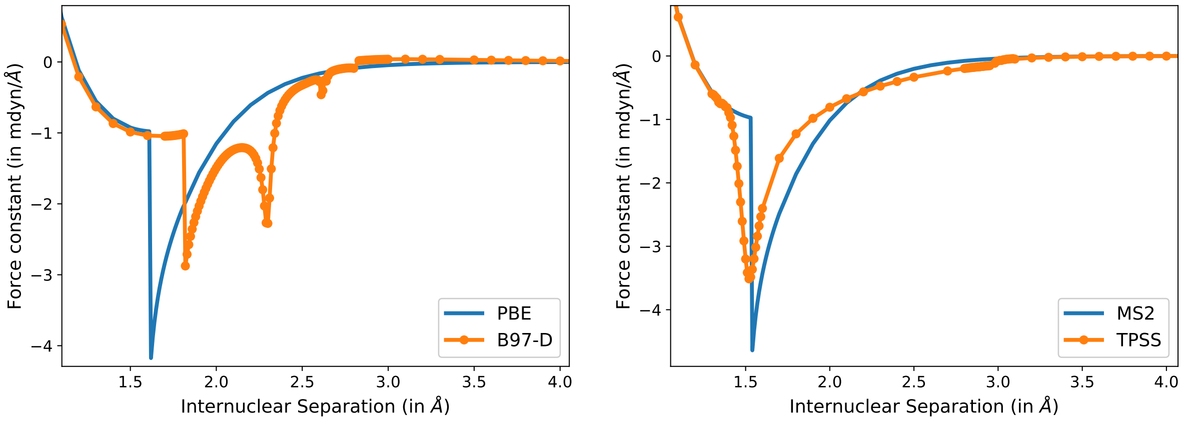


Figure 7.5: Force constant for bond stretching predicted by B97-D (left) and TPSS (right), with an acceptable functional from the same rung of Jacob's ladder as a reference.

The left panel of Fig 7.5 reveals that the B97-D force constant plot has a second peak at around 2.3 Å, which is not a feature present in the exact, HF or PBE/MS2 plots, which corresponds precisely to the point where the dramatic drop of polarizability occurs in the left panel of Fig 7.3. Furthermore, satellite structures are present further along the dissociation curve, which spatially correspond to the same internuclear separation as the secondary unphysical features in the static polarizability curve. It is therefore not unreasonable to conclude that the qualitative errors in the force constant and polarizability stem from the same factors.

The case of TPSS (right panel of Fig 7.5) is more interesting in that there does not appear to be a discontinuity in force constant values at the CF point, unlike HF, PBE, or MS2! A *derivative* discontinuity appears to exist, but the force constant value itself continuously changes with bond stretch. The CF point, and associated discontinuities are ultimately artifacts due to the onset of spin polarization in single determinant approximations that exact quantum mechanics does not predict, and so the absence of a normal CF point is not a sign of unphysical behavior in itself. However, the unphysical polarizability predictions shown in the right panel of Fig 7.3 and the subtle discontinuity in the force constant plot around 3 Å suggest that the lack of a normal CF point is more likely to be a symptom of a problem stemming from ineffective localization of spins than a desirable feature.

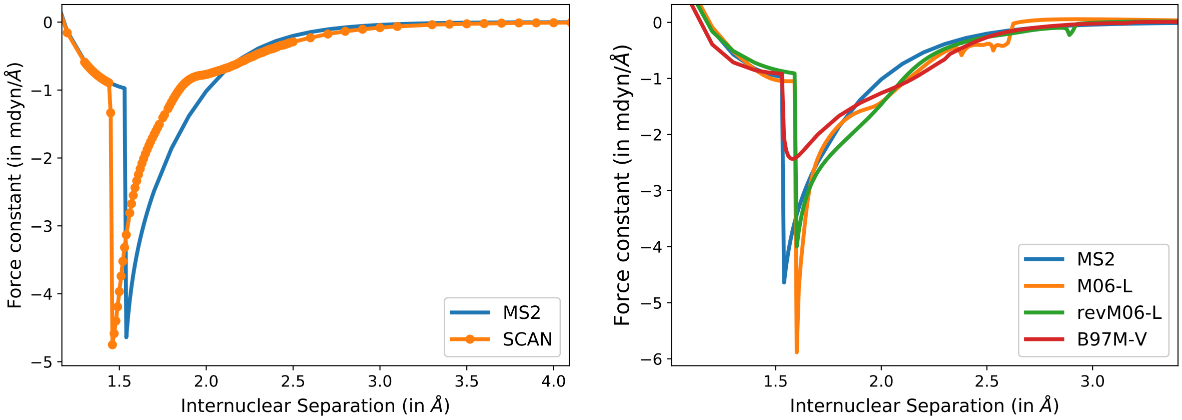


Figure 7.6: Force constant for bond stretching predicted by SCAN (left) and some empirical mGGAs (right), with an acceptable functional from the same rung of Jacob's ladder as a reference.

In contrast, SCAN and the empirical mGGAs M06-L, revM06-L and B97M-V appear to possess a normal discontinuity at their CF points. There are also small discontinuities at 2.6 Å for M06-L and 2.3 Å for B97M-V, corresponding to similar features in the polarizability curve. All three also possess ripple like features further down the dissociation curves. None of them possess any evident points of catastrophic failure (aside from the small discontinuities for M06-L and B97M-V), though their behavior is clearly suboptimal relative to behavior predicted by MS2 (which is qualitatively consistent with exact quantum mechanics, aside from the CF point discontinuity).

7.3.4 Relation to Lowest Eigenvalue of Orbital Rotation Hessian

Eqn 7.3 shows a connection between second order properties (like force constants and polarizabilities) and the inverse of the orbital rotation Hessian matrix. A brief discussion on this Hessian (including its mathematical form and connection to Time-Dependent DFT (TDDFT) for excited states) is supplied in Sec 7.5 for the interested reader, but we will only consider the behavior of the smallest eigenvalue of this Hessian in the main paper. The smallest eigenvalue of the Hessian corresponds to the largest eigenvalue in the Hessian inverse, and consequently could have a disproportionate impact on the second order property predictions. It is therefore interesting to study the behavior of the smallest Hessian eigenvalue over the dissociation curve, and compare/contrast with the observations in the preceding sections.

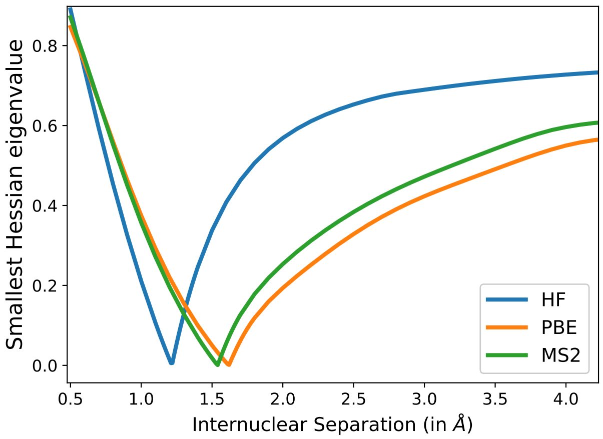


Figure 7.7: Smallest Hessian eigenvalue for HF, PBE and MS2. The eigenvalues asymptotically tend to twice the lowest excitation energy of an H atom, as predicted by each method.

Fig 7.7 shows the smallest Hessian eigenvalue for the three qualitatively acceptable methods HF, PBE and MS2. The behavior is largely as expected, with the eigenvalue decreasing with increasing bond stretch until the CF point, where it becomes zero to permit barrierless transition to the spin polarized state. The effects of the transition manifest themselves in the form of a kink as the two segments of the curve belong to qualitatively different (restricted vs spin polarized) solutions. The eigenvalue increases in magnitude after the CF point due to increasing stability of the spin–polarized solution, ultimately reaching the asymptotic atomic limit.

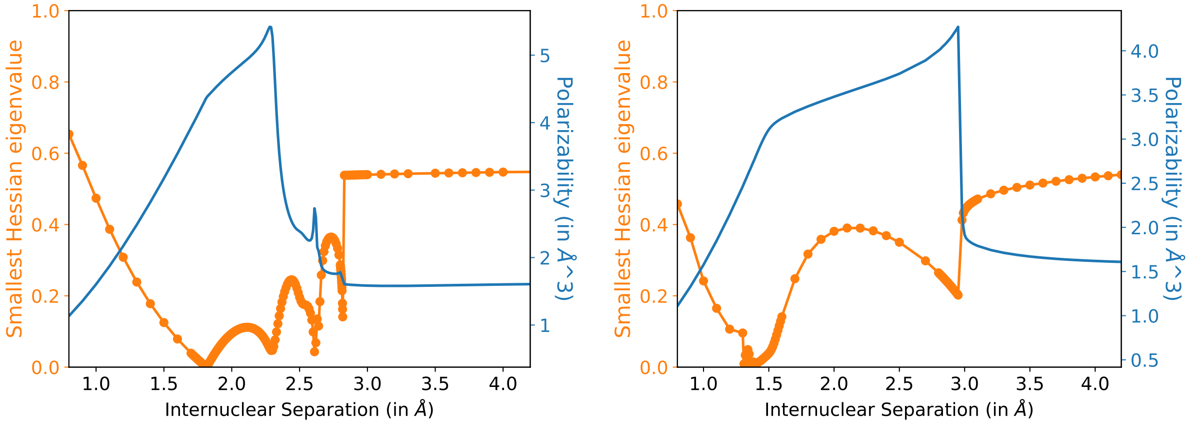


Figure 7.8: Smallest orbital rotation eigenvalues of B97-D (left) and TPSS (right), alongside polarizability predictions.

The problematic functionals, on the other hand, yield less sensible behavior, as can be seen from Figs 7.8 and 7.9. Critically, the unphysical features in the polarizability (and force constant plots) are exactly reflected by the smallest Hessian eigenvalue, indicating that those features emerge from the orbital rotation Hessian inverse term in Eqn 7.3, as opposed to the other terms. Indeed, several features that were somewhat muted in the polarizability plots (like the small discontinuity around $r = 3$ Å in B97-D or the shoulder around 2.5 Å for revM06-L) are considerably magnified by the eigenvalue plot, and are easier to see. In particular, the eigenvalue plot in the left panel of Fig 7.9 clearly reveals how the double peaked structure for the SCAN functional arises from non-monotonicity of the smallest Hessian eigenvalue.

It is however important to recognize that there exists a chance that the eigenvector corresponding to the smallest eigenvalue of the Hessian does not contribute to a specific second order property. Indeed, this is the reason why no discontinuity in the polarizability is observed for nonpolar H_2 at the CF point. It is therefore possible that unphysical orbital Hessian eigenvalues do not correspond to spurious behavior for a *specific* observable, but their role cannot be ignored for *general* second derivative properties. We also note that the terms in the orbital rotation Hessian also play a critical role in calculating excitation energies with TDDFT [251, 277] (see Sec 7.5 for details) and therefore it is near inevitable that the problematic functionals would predict unphysical excited state surfaces. Finally, we observe that the TPSS eigenvalue appears to also go to zero around the CF point, and does not really show unusual behavior immediately beyond it. However, the values appear to be somewhat oscillatory at bond lengths just shorter than the CF point, but this is not really reflected by the observables like the force constant. The origin of the unusual CF point behavior of TPSS therefore remains unresolved with this analysis. Nonetheless, the smallest orbital rotation Hessian eigenvalue proves to be an extremely useful metric in identifying unphysical features in second derivative properties, since it serves as a lodestar by greatly magnifying the errors present in the observables.

7.3.5 Spin Localization

We have hereto conjectured that the unphysical features of the polarizability and force constant plots stem from improper spin density localization into atoms. It would therefore be useful to actually compare the unphysical regions of the plots with a metric for spin density localization. Two such scalar metrics are immediately apparent-the $\langle \hat{S}^2 \rangle$ for the KS determinant, and the overlap between the up and down spin densities. The former however is a problematic metric since none of the observables corresponding to the KS determinant need have any physical meaning aside from the density (which is in principle constrained to be the exact density). The latter on the other hand is quite unambiguous in the case of H_2 as the up and down spin densities should have *zero* overlap in the infinite separation limit (for a single determinant theory), while having perfect overlap till the CF point. Furthermore, the spin densities $\rho_\alpha(\vec{r}) = |\phi_\alpha(\vec{r})|^2$ and $\rho_\beta(\vec{r}) = |\phi_\beta(\vec{r})|^2$ overlap in a manner that is easily calculated from orbital overlap. Specifically, the spatial components of the KS occupied

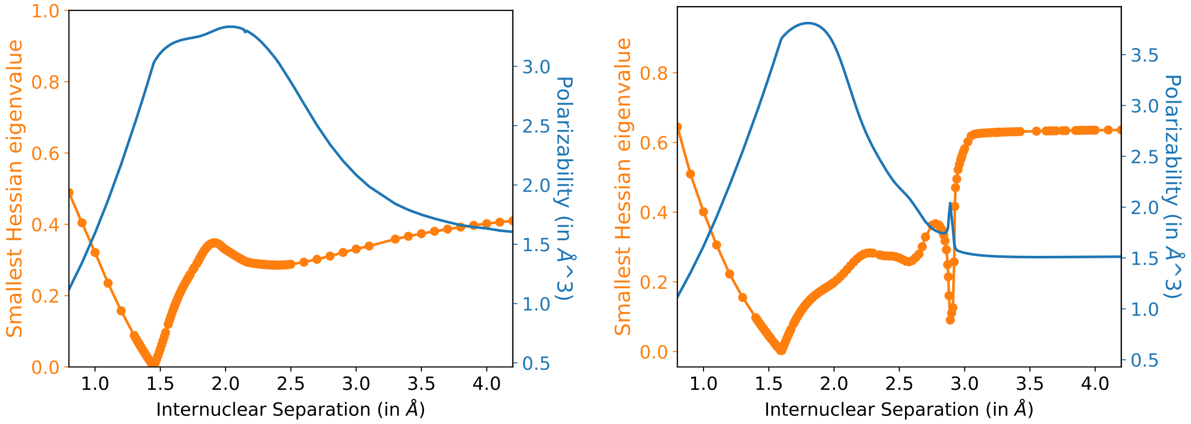


Figure 7.9: Smallest orbital rotation eigenvalues of SCAN (left) and revM06-L (right), alongside polarizability predictions.

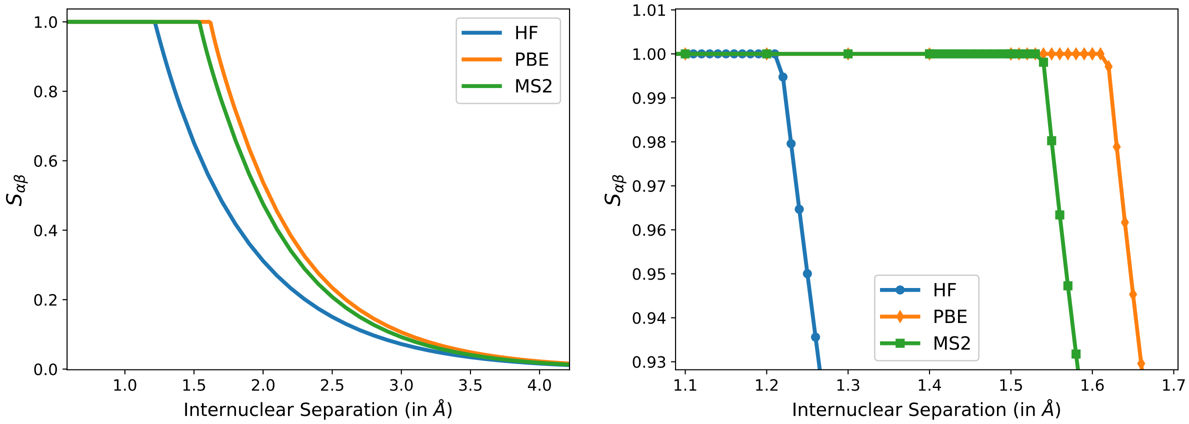


Figure 7.10: Overlap $S_{\alpha\beta}$ predicted by three 'qualitatively acceptable' methods (HF, PBE and MS2) for stretched H_2, highlighting a constant plateau till the CF point, followed by an exponential decay (left). The right panel shows the behavior close to the CF point, highlighting a clear derivative discontinuity for all three methods.

orbitals $\phi_\alpha(\vec{r})$ and $\phi_\beta(\vec{r})$ are nodeless and real for H_2 (on account of being the lowest energy orbitals overall), indicating that the orbital overlap:

$$S_{\alpha\beta} = \left|\int \phi_\alpha(\vec{r})\phi_\beta(\vec{r})d\vec{r}\right| = \int \sqrt{\rho_\alpha(\vec{r})\rho_\beta(\vec{r})}d\vec{r} \tag{7.4}$$

is a convenient measure of the density overlap. It is therefore reasonable to anticipate that $S_{\alpha\beta}$ will be 1 till the CF point (as the spatial orbitals for the two spins will be identical), and would decay exponentially in the asymptotic limit, solely on account of the overlap between the decaying tails of the atomic orbitals. Indeed, our 'qualitatively accurate' methods (HF, PBE and MS2) yield precisely this behavior, as can be seen from Fig 7.10.

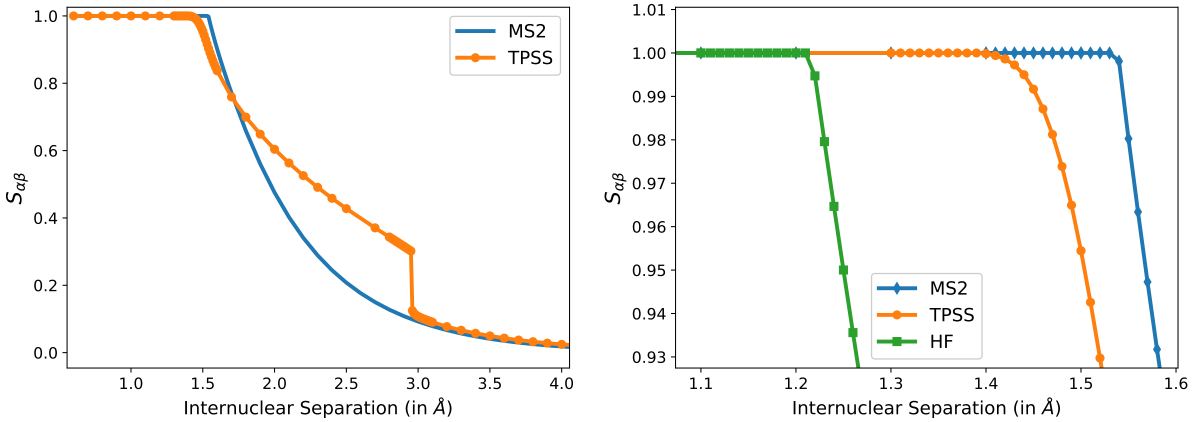


Figure 7.11: Overlap $S_{\alpha\beta}$ generated by TPSS for stretched H_2, with some qualitatively accurate method(s) as reference(s). The left panel shows a clear discontinuity around $r = 3$ Å, which corresponds to discontinuities in the polarizability and force constant seen earlier. The right panel highlights lack of a derivative discontinuity at the CF point, unlike MS2 and HF. This is our strongest evidence showing that TPSS lacks a normal CF point.

We observe much more exotic behavior for some of the 'problematic functionals' encountered earlier. TPSS (Fig 7.11) in particular shows several interesting features, including a discontinuity at 3 Å indicating a discontinuous change in spin density. There is also a lack of an expected derivative discontinuity at the CF point. This seems to suggest that TPSS starts spin polarizing too slowly at the CF point, resulting in a partially spin polarized state till $r = 3$ Å. This incomplete spin polarization consequently preserves a partial covalent bond between the atoms, leading to large polarizabilities of the form seen in Fig 7.3. The slow spin polarization also perhaps leads to $\dfrac{\partial^2 E}{\partial \boldsymbol{\theta} \partial x} = 0$ at the CF point along the zero eigenvalue Hessian mode, voiding the singular term in the Eqn 7.3 and thereby leading to a continuous force constant (unlike all other methods). It further appears that the asymptotic limit of the $S_{\alpha\beta}$ segment up to 3 Å is ≈ 0.14 as opposed to 0 (from fitting to a functional form of $A + Be^{-cx}$), indicating that the state in question corresponds to an entirely different, partially spin polarized UKS state as opposed to the fully polarized state that is the standard solution to the UKS equations beyond the CF point. However, a discontinuous transition around $r = 3$ Å to the fully spin-polarized, bond free state occurs, taking TPSS to the right asymptotic limit of independent atoms. This therefore suggests that there are two potential UKS states predicted by TPSS for stretched H_2, one that is fully spin polarized and one that is partially so, with the latter being energetically preferred at intermediate bond stretch levels.

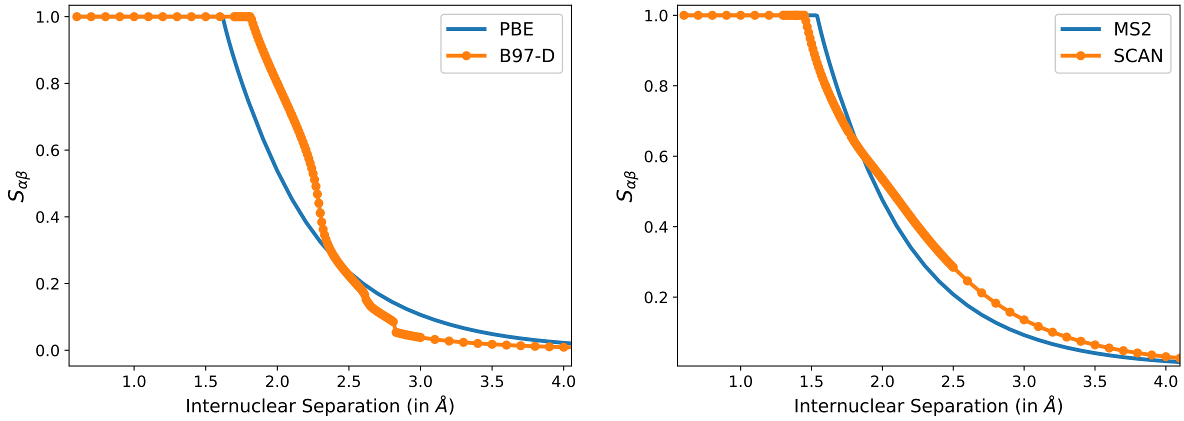


Figure 7.12: Overlap $S_{\alpha\beta}$ generated by B97-D (left) and SCAN (right) for stretched H_2, with some qualitatively accurate method(s) as reference(s).

$S_{\alpha\beta}$ for B97-D similarly shows some unexpected features like changes in the sign of curvature at different points in the curve (Fig 7.12, left panel), a very rapid (and non-exponential) decay around 2.3 Å (which corresponds to the rapid drop in polarizability and the second peak in the force constant), some smaller staircase like features and a very small discontinuity around 2.8 Å. The region of rapid drop in overlap seems to suggest a transition from a partially polarized state with a fractional bond to a fully polarized one, although the continuous nature of the change seems to suggest that it is a single state that smoothly changes its character, unlike the case of TPSS with a discontinuous transition. On the other hand, SCAN has more reasonable behavior (Fig 7.12, right panel), though there seems to be signs of a subtle change in curvature sign around $r = 2$ Å, which matches nicely with the not so subtle behavior polarizability and force constants in that neighborhood, as shown earlier.

Coming to the matter of empirically fitted mGGAs, we find that M06-L has the most problematic behavior, possessing a small ripple (i.e. change in curvature sign) around 2.4 Å, corresponding to a sudden spike in the polarizability prediction, along with small discontinuity at 2.6 Å (see Fig S50 in the supplemental information). The revised revM06-L has fewer objectionable features (see Fig S56 in the supplemental information), with a small kink around 2.8 Å being the most prominent—again corresponding to a spurious peak in the polarizability. B97M-V on the other hand, does not appear to have any visibly obvious unphysical features (see Fig S46, supplemental information), aside from a kink corresponding to the bump in the polarizability at 3 Å. All three functionals however show no unphysicalities as dramatic as the behavior seen for TPSS and B97-D, and on the whole, their performance is mostly satisfactory.

7.3.6 Density Corrected Polarizabilities

The preceding section strongly suggests that unphysical predictions made by the functionals stem from inaccuracies in the underlying density (or more specifically, from the limited extent of spin polarization). It therefore seems instructive to consider the behavior of these functionals when supplied with a qualitatively correct density. We achieve this by calculating static polarizabilities (as predicted from the second order response of the energy to an applied electric field) from HF orbitals, using various functionals. This is essentially the density corrected DFT (DC-DFT) protocol proposed by Kim et al. [278, 279].

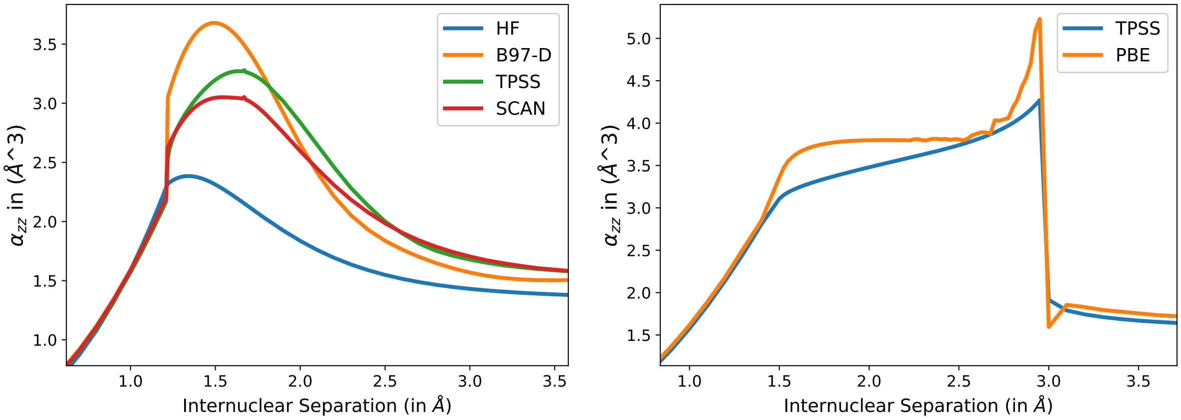


Figure 7.13: Static polarizability parallel to the bond axis generated from HF orbitals (left) and TPSS orbitals (right).

Fig 7.13 (left panel) shows that the unphysical features in the polarizability predictions of B97-D, TPSS and SCAN disappear once the UHF orbitals are used instead of the self-consistent ones. There is however a clear discontinuity at the HF CF point, but this was largely expected due to the presence of terms that did not contribute to Eqn 7.3 on account of self-consistency enforcing the Hellman-Feynman theorem. The behavior otherwise is quite reasonable, and supports the hypothesis that the polarizability errors stem from inaccuracies in the self-consistent density. Conversely, using the TPSS density to predict polarizability causes even the usually well behaved PBE functional to break down (Fig 7.13, right panel) and provide a similar shape as TPSS with self-consistent orbitals. We note that there is no CF point discontinuity when the TPSS density is used (though there is a barely perceptible kink) because TPSS does not have a normal CF point with a derivative discontinuity in $S_{\alpha\beta}$. We also observed much worse breakdown in polarizability predictions (including unphysical negative values!) when HF was combined with TPSS orbitals, or when PBE/HF was combined with B97-D orbitals, indicating that those densities were not suitable for generally well-behaved methods.

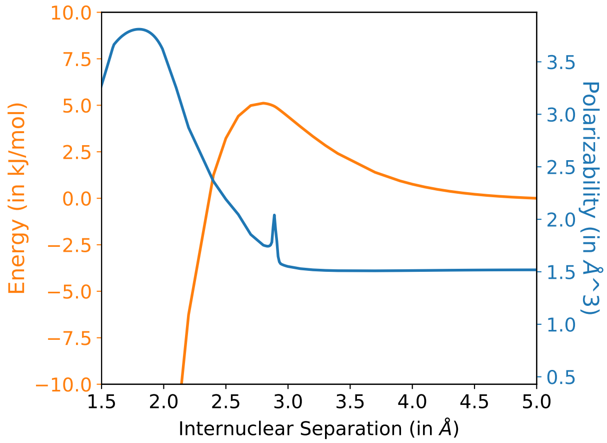


Figure 7.14: Energy of stretched H_2 relative to the dissociation limit, predicted by revM06-L, alongside static polarizability. The unphysical barrier and the strange polarizability spike occur in the same region of the stretch. It is possible that the polarizability spike results from a dipole allowed, low lying excited state, which perhaps mixes into the UKS determinant and spuriously elevates ground state energy.

7.3.7 Hybrid Functionals

Our analysis so far has mostly focused on local functionals (i.e. from the first three rungs of Jacob's ladder), but it is important to consider hybrid functionals as well, due to their widespread use on account of greater accuracy for prediction of energies and properties at equilibrium [55, 80, 208, 226]. Broadly speaking, two classes of hybrid functionals are possible: ones that are derived from mixing existing local functionals with some fraction of HF exchange and ones that are completely parameterized in the presence of HF exchange. Examples of the former are PBE0 (25% HF exchange combined with 75% PBE exchange and 100% PBE correlation), B3LYP, TPSSh and SCAN0, while the latter include the (ω)B97 family of hybrid functionals and M06 [120].

Hybrid functionals of the first type show the same qualitative behavior as the local functional they have been derived from. PBE0 and B3LYP therefore give quite reasonable behavior, producing curves that have the same shape (though not the same values) as the local PBE and BLYP. Conversely, TPSSh and SCAN0 have qualitative failures of exactly the same nature as TPSS and SCAN, suggesting that the observed errors are not delocalization driven (as those errors go down with increasing admixture of HF exchange).

No such reference exists for hybrid functionals of the second type as they were completely parameterized in the presence of HF exchange. Catastrophic failures of similar nature to what was observed for B97-D and TPSS were not observed for M06 [120], B97 [175], B97-2 [118], ωB97X-D [165],ωB97X-V [108] and ωB97M-V [96]. We do however observe small bumps in the 2.5-3 Å interval for polarizability predictions from M06 (Fig S9), B97 (Fig

S5), B97-2 (Fig S2) and ωB97X-D (Fig S20) ωB97M-V has a small discontinuity around 1.95 Å (Fig S19), while ωB97X-V does not have any visually evident problematic features (Fig S21). The bumps and the discontinuity correspond to similar features in the smallest Hessian eigenvalues as well, showing the connection between these features.

7.3.8 Barriers to H Atom Association

There are no local maxima in the energy of a stretched H_2 molecule, as predicted by exact quantum mechanics. Such a maximum would in fact not make physical sense, as it would indicate that there is a barrier to two H atoms associating to form H_2 and so there would be regions where the forces would push two H atoms further apart instead of closer together.

Functional	**Rung**	**Barrier (in kJ/mol)**
B97-D	2	11.7
M06-L	3	13.8
revM06-L	3	5.2
B97M-V	3	1.4
B97-2	4	8.7
B97	4	3.5
M06	4	2.4
ωB97X-D	4	4.0
ωB97X-V	4	0.1
ωB97M-V	4	0.8

Table 7.1: Functionals that predicted a barrier against H+H association, along with barrier values (in kJ/mol) and their rung in Jacob's ladder.

While spurious barriers in dissociation curves are not unknown, they tend to arise mostly from non-variational optimization of parameters in methods like projected coupled cluster [280], the random phase approximation (RPA) [281], approximate adiabatic connection forms [282] or xDH double hybrid functionals [81]. We however find that many density functionals appear to predict an unphysical barrier to association for two H atoms, *despite* self-consistent optimization of orbitals. So far we have noted the presence of a barrier in a number of functionals, which are enumerated in Table 7.1. Fortunately, the widely used HF, PBE, TPSS, SCAN and MS2 remain blissfully barrier free, as do the hybrid methods TPSSh, SCAN0, PBE0, B3LYP and MS2h.

The nearly universal presence of barriers in the B97 type functionals is of particular interest as this highly flexible functional form [175] has been employed to develop a number of functionals that are excellent for prediction of energetics [55, 111]. Many of the more modern B97 variants have barriers that are below 'chemical accuracy' (i.e. 1 kcal/mol$\approx$ 4 kJ/mol) for the specific case of H_2, as can be seen from Table 7.1. However, the presence of such barriers is itself a severe qualitative failure, and there exists a chance that larger spurious barriers will be predicted for more complex species. Future functional development using the B97 form should therefore keep the potential for predicting artificial barriers into account, in the hope that it can be mitigated or at least minimized.

We are not entirely sure if the barriers are necessarily connected to the polarizability catastrophes mentioned earlier. TPSS for instance lacks a barrier, despite having significant issues with polarizabilities. On the other hand, it is certainly possible that fractionally bonded states generated by bond stretches elevate the energy spuriously at intermediate bond stretches due to the presence of high energy ionic contributions. Both revM06-L and M06-L have the energy maximum in the neighborhood of their unphysical polarizability spikes (as can be seen for the former in Fig 7.14), supporting this view. However, the barriers do not disappear when the energy is calculated with HF orbitals instead of self-consistent ones (indeed, they marginally get bigger), indicating that is not just a failure of the self-consistent density. It must however be kept in mind that the use of non self-consistent densities elevates absolute energies, and is likely to have maximum impact at points where the self-consistent density is most different from the supplied density, which are the problematic points along the dissociation curve. So, it is perhaps expected that the barriers will persist even after density correction. Furthermore, it is known that errors in forces can be directly connected to errors in the electron density for H_2 via the Feynman electrostatic theorem [283], indicating that it is quite possible that poor spin localization could be responsible for the spurious forces that generate the barriers). Nonetheless, we are not quite sure of the exact reasons behind the origin of these barriers at present, but would recommend caution in applying the problematic functionals for ab-initio dynamics due to the potential of obtaining spurious forces and unphysical barriers.

7.4 Discussion and Conclusions

KS-DFT methods are typically accurate and efficient for treating electronic structure in the absence of strong correlation effects. Prototypical examples include closed-shell molecules and their separated, spin-polarized, radical fragments. Yet in applications like *ab initio* dynamics, such bonds may be made or broken, and density functionals may therefore have to cope with the strongly correlated spin-recoupling regime. The closed shell bonded regime and the spin-polarized broken bond regime are joined via the Coulson-Fischer (CF) point, which signals the onset of strong correlations that cause spin polarization for approximate single reference methods. High accuracy in the recoupling regime is not readily available today, but it is reasonable to expect that UKS models should smoothly join the CF point to the broken bond regime, similar to UHF. Such functionals can be considered to be well-behaved.

In this work, we have examined the homolytic bond-breaking of the H-H bond as a simple prototype, using second derivative properties (polarizabilities and force constants) at stretched bond lengths as sensitive probes of whether or not functionals are well behaved.

We consequently find that the reasonable expectation that UKS models can describe the dissociation of H_2 (a nonpolar, singly bonded neutral molecule) with qualitative accuracy to be flawed. A number of popular (PBE, B3LYP) and recent (MS2) functionals can in fact describe properties with similar qualitative features as exact quantum mechanics, but their success is by no means indicative of the general UKS performance. A fair number of functionals (both empirically trained and non-empirically constrained) appear to yield unphysical predictions for static polarizability and force constant for bond stretching. The most egregious offenders are the empirically fitted B97-D GGA and the nonempirical TPSS meta-GGA, which appear to predict a fractionally bonded state with incomplete spin localization and high polarizability at intermediate bond stretch levels, and subsequently undergo a dramatic transition to the atomic asymptotic limit with localized spins. Other functionals like SCAN and M06-L have fewer objectionable features, but their predictions do not appear physical at all times. All these unphysicalities appear to originate from the smallest eigenvalue of the orbital Hessian matrix, whose erratic behavior in turn appears to stem from incorrect levels of spin polarization in the self-consistent density. Consequently, the polarizability errors can be corrected via the use of physically correct, completely spin polarized densities such as those from UHF, indicating that the errors here are essentially 'density driven'.

We also find that a large number of functionals predict a spurious barrier impeding the association of free H atoms to form H_2, although a clear connection between these barriers and the polarizability catastrophes cannot yet be clearly drawn. It is however quite clear that the presence of such barriers would be problematic for any potential application of these functionals in regimes involving bond breaking. In particular, *ab initio* molecular dynamics simulations using these methods would suffer from incorrect forces and artificial barriers, potentially affecting the conclusions. We therefore recommend extensive investigation of the potential energy surface along the reaction coordinate before employing these functionals for dynamical studies, in order to ensure that unphysical effects are kept to a minimum.

The errors we have found appear to be neither delocalization nor strong correlation (as HF and several other functionals do not encounter these issues). We therefore believe that we have identified a new class of errors that affects some (perhaps many) approximate density functionals, and can strongly manifest itself in the dissociation of nonpolar bonds in neutral molecules like H_2. These errors appear to mostly stem from inaccuracies in the underlying self-consistent density than an obvious defect of the functional form itself. Further work is required to characterize this class of errors, in order to mitigate their occurrence in future functionals.

7.5 Appendix: Orbital Rotation Hessian

Sec III D has a discussion about the Hessian of the electronic energy against orbital rotations and how it is useful for characterizing failures in predicting static polarizabilities and force constants. The general idea behind this Hessian dates back to the pioneering work by

Thouless on the stability of HF solutions [277]. A key result is that it is possible to rotate between two non-orthogonal Slater determinants via an exponential of single excitations. Mathematically, this means:

$$|\Psi\rangle = e^{\sum\limits_{ia}\left(\theta_{ia}a_a^{\dagger}a_i - \theta_{ia}^{*}a_i^{\dagger}a_a\right)}|\Phi\rangle \tag{7.5}$$

where $|\Psi\rangle$ and $|\Phi\rangle$ are two non-orthogonal Slater determinants, while i and a are indices for occupied and virtual orbitals corresponding to $|\Phi\rangle$, respectively. The orbital degrees of freedom are represented by θ_{ia}, which can be varied to find stationary points of the energy of a Kohn Sham (or Hartree-Fock) single determinant energy, E. A stationary point has zero energy gradient with respect to all θ_{ia} such that the vector, $\dfrac{\partial E}{\partial \boldsymbol{\theta}} = 0$. A stationary determinant $|\Psi\rangle$ however can only be considered to be stable if the Hessian of second derivatives of E with respect to $\boldsymbol{\theta}$ (i.e. the matrix $\mathbf{H} = \dfrac{\partial^2 E}{\partial \boldsymbol{\theta}^2}$) is positive semidefinite. Otherwise $|\Psi\rangle$ would only represent some (possibly high order) saddle point in orbital space, and not a true (local) minimum. Eqn 7.5 can be employed to find the general form of the Hessian (subject to given constraints on orbitals), and many important resulting stability conditions (restricted to unrestricted orbitals, real to complex orbitals etc) have been enumerated for Hartree Fock theory [15]. Generalization to Kohn-Sham DFT is quite straightforward [251].

We are however only interested in stable, unrestricted solutions with real orbitals. This causes the Hessian to readily simplify to be $\mathbf{H} = 2(\mathbf{A}+\mathbf{B})$, where the $\mathbf{A}$ and $\mathbf{B}$ matrices are defined as follows:

$$A_{ia,jb} = (\epsilon_a - \epsilon_j)\,\delta_{ij}\delta_{ab} + \langle ib|\,|aj\rangle \tag{7.6}$$

$$B_{ia,jb} = \langle ij|\,|ab\rangle \tag{7.7}$$

in terms of occupied-virtual orbital pairs ia and jb. $\epsilon_{\{a,i\}}$ are orbital energies and $\langle ij|\,|ab\rangle$ are antisymmetrized two-electron repulsion matrix elements in physicist's notation [14]. These Hartree-Fock based expressions for $\mathbf{A}$ and $\mathbf{B}$ can be readily generalized to UKS [251, 284], which also has a Hessian that is represented by $2(\mathbf{A}+\mathbf{B})$. The $\mathbf{A}$ and $\mathbf{B}$ matrices furthermore are employed in the random phase approximation (RPA) [277] and linear-response methods for excited states. Specifically, Time-dependent DFT (TDDFT) for excited states involves solving for positive eigenvalues of $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B} & -\mathbf{A} \end{pmatrix}$, highlighting the intimate connection between standard excited state methods and the orbital rotation Hessian [251].

7.6 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: figures of polarizabilities, smallest Hessian eigenvalues, force constants and $S_{\alpha\beta}$ (pdf), and raw data (xlsx).

Chapter 8

Afterword to Electrical Response Properties

Now far ahead the Road has gone,
Let others follow it who can!
Let them a journey new begin,
But I at last with weary feet
Will turn towards the lighted inn,
My evening-rest and sleep to meet.
J. R. R. Tolkien "The Road Goes Ever On"

This half of the dissertation describes the ability of KS-DFT to predict electrical response properties like dipole moments, second cumulants of the electron density and the static dipole polarizability. In particular, CCSD(T)/CBS level benchmarks at molecular equilibrium geometries for 100 or more chemical systems were presented for each of the three properties, in order to assist training and assessment of DFAs. The performance of modern DFAs, as well as several classics were also assessed over the datasets, to gauge how well the current state of the art can model the electrical response of molecules. The results indicate that some of the most successful modern mGGA based functionals for energy predictions do not show comparable accuracy for response properties, being uninspiring that regard. Consequently, there remains considerable ground to cover with regards to developing functionals that accurately predict both energies and response properties. In addition, we utilize qualitative failures in predicting these properties at stretched bond lengths, to diagnose shortcomings with certain DFAs. In total, it is hoped that the material presented will enable the development of functionals with reasonable predictive power across a range of applied external potentials, or at least caution against the use of existing DFAs that are found to be problematic.

Has it been successful in this regard? The dipole and static polarizability studies have proven to be reasonably influential at present, at least in terms of citations gathered. The second cumulant investigation has not had enough time to mature, while the polarizability of stretched H_2 segment has mostly been utilized as a cautionary note in some studies, indicating that it has not quite broken out of the niche. Both the dipole and polarizability assessments have seen extensive use in justifying choice of DFA for computing density related and electrical response properties. As such, they appear to have conveyed to the community what approaches are most suitable for such calculations, as well as what levels of error might be expected [285–293]. In particular, they have been used to assist parameterization of classical charge distributions [286, 290–292] for use in force-fields. The dipoles and (to a lesser extent) the polarizabilities have also been employed for assessing DFAs [114, 224, 225, 294], wavefunction methods [221, 295] and even a semiempirical method [296] at the time of development or soon after. This indicates the increasing acceptance within the scientific community that density and response properties ought to also be considered while gauging the overall efficacy of QM approximations. Perhaps the best reflection of this are recent efforts to generate benchmark excited state dipole datasets, for assessment of more approximate techniques [297, 298]. However, response information have not yet directly been utilized as training sets, to the best of our knowledge.

Would we have done anything differently, with the benefit of hindsight? The natural next step would be to incorporate post CCSD(T) corrections to the dataset, to bring them closer to FCI. However, this is unlikely to be a major issue for even relatively multireference species like O_3 as KS-DFT errors are likely to far exceed any such correction in magnitude. In a similar vein, use of explicitly multireference methods like selected CI for the stretched bond dipole studies could be of interest, especially in light of reported differences between CCSD(2) and MRCISD+Q for stretched HF [299]. In addition, it would have been useful to investigate finite difference errors in greater depth for the equilibrium polarizability dataset. The analysis for acceptable field strengths was done at the HF level, which is sufficient for wavefunction methods, as well as for double hybrids and (likely) hybrids. However, local functionals often predict spuriously large hyperpolarizabilities [260] that can lead to larger finite difference errors for polarizabilities with such functionals. Indeed, Ref 238 reports that the reported PBE polarizabilities for six species (Li, FH-OH, HO_2, NaCl, NaCN, and BeH) in our polarizability dataset had some contamination from higher order hyperpolarizabilities. However, exclusion of these molecules did not affect any of the conclusions of Chapter 6, and the data still reflects the utility of DFAs for predicting the response of the energy to applied fields. The benchmark quality was also not compromised as it was the CCSD(T) level, where the spurious hyperpolarizability problem should not exist.

In conclusion, it appears that density and response based properties are likely to see significant use in development and assessment of new quantum chemistry methods in future. This will require the use of benchmark datasets like this one and others [297, 300], with many more likely to come. It would be interesting to see if such data would constrain DFA development to hew closer to the ‘true path’. In general, arbitrary response property predictions expose DFAs to the full range of local potentials permitted by the HK theorems and in principle allow for all ground state possibilities to be explored. Additional data, if necessary, would have to originate from electronic excited states, as the extremal densities of the exact functional correspond to such states. Further discussion about DFT excited states is provided in the subsequent chapters of this dissertation.

Part II

Electronic Excited States

Chapter 9

Foreword to Electronic Excited States

A ship then new they built for him
of mithril and of elven-glass
with shining prow; no shaven oar
nor sail she bore on silver mast:
the Silmaril as lantern light...
and wings immortal made for him,
and laid on him undying doom,
to sail the shoreless skies and come
behind the Sun and light of Moon.
J. R. R. Tolkien "Song of Eärendil"

Electronic excited states refer to solutions of the electronic Hamiltonian $\mathbf{H}_e$ that are not the lowest energy ('ground') state. These states often arise via light induced excitation from the ground state, and consequently play a key role in the photophysics and photochemistry of chemical systems [301]. Characterization of such states is consequently not only of interest from a basic science perspective, but is also critical for efficient design of photovoltaic materials, photocatalysts, lighting devices etc. The development of theoretical methods to model electronic excitations is thus of considerable importance. Progress in these directions has been somewhat slower than comparable efforts to model the ground state, but new and exciting developments in the area have come a long way in bridging the gap.

Very accurate wavefunction methods are extremely useful for reliably getting accurate excitation energies and properties. However, the computational expense of such methods mostly restrict their use to the development of benchmark datasets [297, 302] (against which more approximate methods of lower computational complexity can be assessed). The success of KS-DFT in efficiently modeling electronic ground states appears to suggest that DFT protocols can also be used to affordably access electronic excitations with a reasonable level of accuracy [251, 303–305].

9.1 Linear Response Time-Dependent DFT

The most widely used excited state protocol is likely linear-response time-dependent DFT [251, 306] (LR-TDDFT, henceforth refered to as TDDFT). TDDFT obtains excited state energies and properties via the linear-response of a ground state DFT solution to time-dependent electric fields. A full derivation of the TDDFT equations is well described in Ref. 251. We therefore only briefly summarize the key results herein.

Let the ground state KS determinant have occupied spin orbitals $\phi_{\{i,j,k,l\ldots\}}$ and virtual (unoccupied) spin orbitals $\phi_{\{a,b,c,d\ldots\}}$, that arise from a general exchange-correlation functional E_{xc}. E_{xc} can contain orbital dependent terms like HF exchange, as well as contributions from the electron number density $\rho(\vec{r})$. In the generalized KS framework [307], TDHF [37, 308] is the special case where E_{xc} contains *only* 100% HF exchange. Furthermore, let us assume all orbitals are real valued. The TDDFT excitation energies $\{\omega_k\}$ are then found via solving the following non-Hermitian generalized eigenvalue problem:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega_k \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{9.1}$$

where the $\mathbf{A}$ and $\mathbf{B}$ matrices are:

$$A_{ia,jb} = (\epsilon_a - \epsilon_i)\,\delta_{ij}\delta_{ab} + \langle ij|ab\rangle + \langle ij|\, f_{xc}\, |ab\rangle \tag{9.2}$$

$$B_{ia,jb} = \langle ib|aj\rangle + \langle ib|\, f_{xc}\, |aj\rangle \tag{9.3}$$

ϵ_p is the energy of orbital ϕ_p and the two electron integrals $\langle ij|ab\rangle$ and $\langle ij|\, f_{xc}\, |ab\rangle$ are :

$$\langle ij|ab\rangle = \int \phi_i(\vec{r}_1, s_1)\,\phi_j(\vec{r}_2, s_2)\,\frac{1}{|\vec{r}_1 - \vec{r}_2|}\phi_a(\vec{r}_1, s_1)\,\phi_b(\vec{r}_2, s_2)\,d\vec{r}_1 d\vec{r}_2 ds_1 ds_2 \tag{9.4}$$

$$\langle ij|\, f_{xc}\, |ab\rangle = \int \phi_i(\vec{r}_1, s_1)\,\phi_j(\vec{r}_2, s_2)\,\frac{\delta^2 E_{xc}}{\delta\rho(\vec{r}_1)\,\delta\rho(\vec{r}_2)}\phi_a(\vec{r}_1, s_1)\,\phi_b(\vec{r}_2, s_2)\,d\vec{r}_1 d\vec{r}_2 ds_1 ds_2 \tag{9.5}$$

We also note that a similar formalism is utilized by the Bethe-Salpeter equation [309] for electronic excitations, which utilizes a Greens function based approach to perturbation theory. Such techniques are experiencing increased use for molecular systems [310–312].

Eqn. 9.1 is closely related to the stability of the reference KS solution to orbital rotations. An unrestricted KS (UKS) solution with real valued orbitals is stable if both $\mathbf{A} + \mathbf{B}$ and $\mathbf{A} - \mathbf{B}$ are positive semidefinite [15]. In particular, a restricted KS (RKS) solution is stable against spin polarization if $\mathbf{A} + \mathbf{B}$ is positive semidefinite. If not, a spin-polarized UKS solution of lower energy exists (see also Chapter 7). The connection between stability and Eqn 9.1 becomes clear when the latter is simplified to:

$$(\mathbf{A} - \mathbf{B})(\mathbf{A} + \mathbf{B})(\mathbf{X} + \mathbf{Y}) = \omega^2 (\mathbf{X} + \mathbf{Y}) \tag{9.6}$$

$\mathbf{A}$ is typically much larger than $\mathbf{B}$, since the former contains orbital energy differences (that has mean-field one body contributions) that should be much larger than the purely two body terms (which are the sole constituents of $\mathbf{B}$) for KS-DFT to be viable for the ground state [251]. An instability therefore is an indication of issues with the RKS ground state solution, and often reflects the presence of strong correlation.

The non-Hermitian nature of Eqn. 9.1 leads to the possibility of complex TDDFT eigenvalues if the stability conditions are violated. Partly for this reason, it has been suggested that setting $\mathbf{B} = 0$ should be a useful approximation. The resulting eigenvalue equation is simply $\mathbf{AX} = \omega\mathbf{X}$, and is called the Tamm-Dancoff approximation (TDA) [313–315]. This is typically cheaper than full TDDFT and has the added benefit of only requiring diagonalization of a Hermitian matrix, precluding any possibility of unphysical imaginary excitation energies. TDDFT and TDA are also formally size-consistent [251], making them appealing for studying PESs. Indeed, it has been suggested that TDA is arguably more reliable than TDDFT for explorations of PESs [316]! A fundamental examination of the effect of orbital instabilities on TDDFT and TDA excited states will be presented later in Chapter 10.

We note that both TDDFT and TDA have some rather desirable features. A large number of excited state solutions can be simultaneously obtained via solving the associated eigenproblems, and so it is possible to calculate all states within a desired range of energies. Consequently, no prior knowledge about the nature of target states is generally needed. The protocols are also quite computationally affordable, if Davidson type solvers [317, 318] are utilized. The computational efficiency stems from the ease of evaluating matrix-vector products of the form $\mathbf{AX}$, which roughly has the same cost as a single ground state KS-DFT iteration [319]. TDDFT or TDA calculations therefore have the same practical scaling as ground state KS-DFT.

The special case of E_{xc} being purely the HF exchange functional for TDDFT and TDA merits particular attention. In this limit, both become pure wavefunction methods that can exist independent of KS-DFT. TDHF excitation energies in particular are connected to the correlation energy within the random phase approximation (RPA) [251]. TDHF/TDA turns out to be a configuration interaction (CI) method as:

$$A_{ia,jb} = \langle \Phi_i^a | \mathbf{H}_e \left| \Phi_j^b \right\rangle - E_{HF} \tag{9.7}$$

i.e. $A_{ia,jb}$ is the matrix element of the electronic Hamiltonian $\mathbf{H}_e$ between two singly excited determinants $\left|\Phi_j^b\right\rangle$ and $|\Phi_i^a\rangle$, minus the energy E_{HF} of the HF ground state reference $|\Phi\rangle$. Since $\langle \Phi_i^a | \mathbf{H} | \Phi \rangle = 0$ from Brillouin's theorem [14], we have

$$\mathbf{H} = \begin{pmatrix} E_{HF} & 0 \\ 0 & \mathbf{A} + E_{HF}\mathbf{1} \end{pmatrix} \tag{9.8}$$

within the Hilbert space spanned by the reference determinant and all single excitations $\{|\Phi_i^a\rangle\}$. Consequently, the TDHF/TDA excitation energies are differences between E_{HF} and other eigenvalues of this Hamiltonian, which is exactly configuration interaction with all single (CIS) substitutions from the HF determinant [320]. CIS is therefore a simple and well-behaved member of the TDDFT family of methods, albeit one only suitable for singly excited states. The density overlocalizing tendency of ground state HF however prevents it from attaining quantitative accuracy even for single excitations, leading to errors on the order of 1-2 eV, [251] especially in the form of systematic overestimation for charge transfer (CT) excitation energies [251, 321]. CIS is also not always spin-pure for open-shell systems as some essential doubly excited states must be included, but this can be addressed via methods like extended CIS (XCIS [322]).

The behavior of CIS is in fact quite instructive for comprehending the limitations of general TDDFT. TDDFT is formally exact [250] if the exact time-dependent exchange-correlation (xc) functional is employed. In practice, approximate, time-independent ground state xc functionals are instead utilized, (the so called adiabatic local-density approximation/ALDA [251]). This route is capable of yielding reasonable results [319, 323, 324] for low lying valence excitations in single reference species (see Ref 319 for a recent benchmark of TDDFT excitation energies). However, ALDA restricts TDDFT to single excitations alone, in close analogy to CIS. This makes it impossible to model doubly (or higher) excited states [325–327]. In Chapter 10, we will also show that TDDFT methods are incapable of describing bond dissociations in the excited state.

Furthermore, ALDA only permits orbital relaxation to LR (i.e. first order), which is problematic for modeling excited states that have substantially different densities than the ground state. The classic example is long range CT, for which TDDFT excitation energies are strongly dependent on the fraction of HF exchange present in the functional [251]. Similar behavior is also observed for Rydberg states [325] and core excitations [328] (effectively CT out of core orbitals). The CT problem is largely a consequence of ground state delocalization error [64, 95]. However, the LR protocol magnifies this error to catastrophic proportions in a manner that is atypical for ground state calculations. Let us consider the lowest energy CT excitation between an electron donor D and an electron acceptor C, at infinite separation. TDDFT predicts the excitation energy to be the difference in energy between the LUMO of C and HOMO of D [251], which proves to be quite inaccurate in practice and is also quite functional sensitive (being underestimated by local functionals and overestimated by pure HF). However, we know that the true excitation energy should equal the difference of the ionization potential (IP) of D and the electron affinity (EA) of C. Both are typically well approximated by ground state KS-DFT [55, 111, 329], indicating the LR protocol is the principal problem. We consider CT from NH_3 to an F_2 molecule 1000 Å away as an example. Fig 9.1 shows that TDDFT predictions as a function of functional span a wide range from ~ 0 (TD-LSDA) to 13 eV (TDHF), as shown in Fig 9.1. It is worth noting that optimal tuning of density functionals [330] (via enforcement of Koopman's theorem [14, 331]) can assist in better modeling of long range CT, but entails system-specific functional optimization and might not be as effective at intermediate separations [332].

9.2 Orbital Optimized DFT

Rather than using TDDFT to evaluate the preceding long range CT excited state, we could instead consider modeling it as a supersystem consisting of NH_3^+ and F_2^- fragments. This

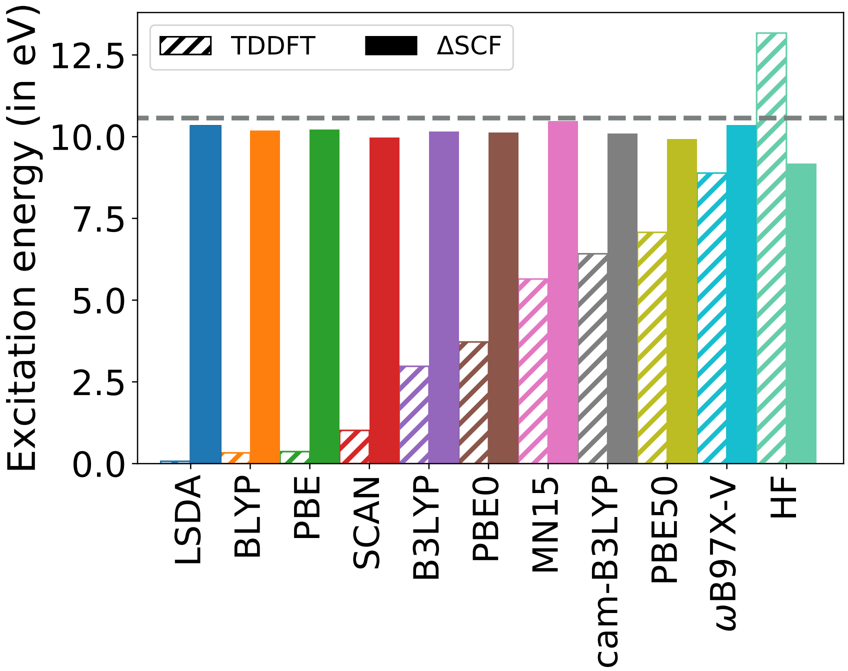


Figure 9.1: Comparison of TDDFT and ΔSCF for the lowest CT triplet excited state of NH_3——F_2 at 1000Å separation (between the N and closest F atom), with various functionals and the def2-QZVPPD basis. The dark line represents the estimate obtained from vertical IP of NH_3 and vertical EA of F_2, evaluated via CCSD(T) [23] at the complete basis set (CBS) limit. The electrostatic interaction between unit point charges at 1000 Å separation is a negligible $\sim$ 0.015 eV.

can be achieved by separately optimizing the orbitals of the charged fragments, constructing a guess density from these fragment densities [333] and subsequent relaxation of the supersystem orbitals to the closest stationary point to this initial guess (which ought to preserve integer charges on the fragments). The excitation energy can then be computed as the energy difference between this self-consistent field (SCF) solution and the ground state SCF solution, leading to this approach being termed as ΔSCF. Fig 9.1 shows that this orbital optimized (OO) approach yields much more reasonable results than TDDFT with the same functionals, and has much lower functional sensitivity. This indicates that state specific OO-DFT could be effective in addressing several of the challenges faced by TDDFT. Indeed, such methods precede TDDFT in the literature [264, 334–337] (dating back to at least Phllipson and Mulliken's 1958 work on the $^3\Sigma_u^+$ and $^1\Sigma_u^+$ states of H_2 [334] with HF), but were not widely used after modern TDDFT implementations came into existence. In our opinion, there are three principal reasons for this:

1. OO methods necessarily require information about the nature of target states, and are

thus not very 'black box'. For instance, we had to explicitly specify an initial guess of NH_3^+ and F_2^- for the example shown in Fig 9.1. In contrast, TDDFT is capable of simultaneously computing multiple states without any prior knowledge about the nature of individual states.

2. Excited states are typically saddle points rather than minima of the energy in orbital space.[1] Optimization of excited state specific orbitals is thus challenging, as it is easy to 'slip' into a nearby local minimum instead of the desired state, which is often described as *variational collapse*. As an example, the lowest triplet state of the NH_3—F_2 supersystem is a local excitation on F_2 ($\sim$ 3 eV excitation energy), and routine SCF cycles would typically land on this state instead of the desired CT state with $\sim$ 10 eV excitation energy.

3. Excited states are often intrinsically multireference. KS-DFT is a single determinant theory by construction, and it is not straightforward to account for multiconfigurational effects. This is generally not an issue for TDDFT out of closed-shell determinants, as the LR protocol gives a formally appropriate route for 'coupling' multiple excited configurations, in a manner analogous to CIS. It is nonetheless worth noting that the lack of double excitations prevents open-shell TDDFT from fully addressing multiconfigurational effects, as no rigorous analogues to XCIS [322] exist.

The first challenge is unavoidable in some regards, as some information about the target state is needed to provide an initial guess for orbital optimization. Specifically, it is extremely useful to know the potential electronic configurations the target state could have, out of the combinatorially scaling possibilities in the full Hilbert space. An initial TDDFT calculation could in fact be useful in identifying the electronic configuration for singly excited states, that could be subsequently used to initiate OO calculations. Doubly (or higher) excited states are more challenging as they are inaccessible in TDDFT, and more rigorous multireference wavefunction methods capable of identifying important configurations are unlikely to be affordable for sizeable systems. A combination of chemical intuition and orbital energies could be effective in generating potential guess configurations for such excited states, if the precise nature of the desired states is not known in advance. We will consequently focus on the variational collapse and multiconfigurational state problems.

---

[1] This can be intuitively understood by noting that the Hessian of the energy vs orbital rotations for a single, spin-unrestricted determinant is $2\left(\mathbf{A}+\mathbf{B}\right)$. Considering only mean-field one body terms (i.e. $B_{ia,jb} \approx 0$ and $A_{ia,jb} \approx \left(\epsilon_a - \epsilon_i\right)\delta_{ij}\delta_{ab}$), we find that the Hessian is diagonal, with the eigenvalues being differences in energy between unoccupied and occupied orbitals. Normally, there exists at least one excited state occupied i orbital with higher energy ϵ_i than an unoccupied orbital a with energy ϵ_a (i.e. non-Aufbau filling generally leads to at least one pair i, a for which $\epsilon_a \leq \epsilon_i$). The Hessian therefore has at least one negative eigenvalue within this approximation, corresponding to at least a first order saddle point.

9.2.1 Variational Collapse

The variational collapse problem has long been a major barrier to widespread use of OO methods [325, 338] beyond the lowest energy state within each spin manifold (such as the lowest energy triplet). It is possible to avert this for certain problems via application of constraints, such as the lowest energy long-range CT excitation shown in Fig 9.1, where individual charges on fragments are unambiguous and can be constrained [339]. Similarly, a specific non-Aufbau configuration can be enforced over all SCF cycles when the orbital energy ordering remains unchanged (such as ensuring a 2s hole for an Ar ion [340]). However, not all problems are as clear cut—necessitating development of more general solutions. One of the most widely used approaches is the maximum overlap method (MOM) [341], which is used in conjunction with Fock matrix (**F**) diagonalization based methods like DIIS [342]. Normally, the selection of occupied orbitals after each SCF cycle is done on the basis of energy, such that the lowest energy levels are filled first (Aufbau principle). MOM instead selects occupied orbitals via maximizing the overlap between the newly constructed determinant with the determinant from the previous iteration, thereby preventing any dramatic change and permitting smooth relaxation of an initial non-Aufbau configuration to an extrema.

Although effective in many cases, MOM cannot always prevent variational collapse as orbitals can continuously change character back to the ground state over multiple steps [343, 344]. This led to the development of the Initial MOM (IMOM) method [344], where the overlap was now maximized relative to the initial set of orbitals to prevent continuous drift down to the ground state. However, dramatic changes in electronic configuration are possible in IMOM when multiple selections can lead to similar overlaps, leading to potential oscillatory behavior and convergence failure [345]. An alternative route to avoiding variational collapse within the repeated **F** diagonalization scheme is through the use of level shifts [346]. This entails shifting the unoccupied orbitals up in energy prior to **F** diagonalization, which both permits Aufbau filling of orbitals (as the undesired levels are pushed up and thus unlikely to be filled) and decelerates occupied-virtual mixing, permitting slow but steady convergence.

Repeated **F** diagonalization based approaches however do not guarantee convergence, as is often painfully evident for nontrivial ground state computations. The most robust solvers for these problems are (quasi-) Newton schemes [347] like geometric direct minimization (GDM) [348], that explicitly attempt to minimize the energy and thus guarantee descent, step by step. Such minimizers however are definitionally unsuited for saddle point convergence, indicating that it might be useful to transform the energy extremization problem as a minimization scheme for some other function. A natural choice in this regards is the variance ($\boldsymbol{\sigma}^2$) of $\mathbf{H}_e$ (i.e. $\boldsymbol{\sigma}^2 = \mathbf{H}_e^2 - \langle \mathbf{H}_e \rangle^2$), as every energy eigenstate corresponds to a global minimum of $\langle \boldsymbol{\sigma}^2 \rangle$ [349, 350]. Recent work by the Neuscamman [351, 352] and Van Voorhis [353, 354] groups have examined $\mathbf{H}_e^2$ based approaches, obtaining very promising results. However, $\mathbf{H}_e^2$ is a rather difficult quantity to work with, as it contains four-particle operators. Furthermore, there is no analogue of $\mathbf{H}_e^2$ in DFT (i.e. no functionals for $\langle \mathbf{H}_e^2 \rangle$). With these motivations in mind, we will present an alternate strategy for excited state orbital optimization through use of quasi-Newton minimizers, in Chapter 11.

9.2.2 Multiconfigurational States

Excitation of one or more electrons tends to break existing spin pairs in chemical systems, which is generally the origin of multiconfigurational character in excited states. It must however be noted that not all excited states necessarily require multiple determinants for a proper representation. States in which all unpaired electrons have the same spin can in fact be represented by single determinants, with well known examples being low lying triplet states of closed shell species (like the $1\mathrm{s}^1 2\mathrm{p}_z^1$ triplet state of He or any other triplet with only two unpaired electrons), certain double excitations [355] (such as $\mathrm{n}^2 \rightarrow (\pi^*)^2$ in HCHO) and single excitations to/from singly occupied levels in open-shell systems [356]. DFT optimization of a single Slater determinant is thus sufficient for such states, and the resulting protocol is termed as ΔSCF. ΔSCF is also used to describe ionization from orbitals other than the HOMO, as they lead to formation of excited states of the cation [340].

Singlet single excitations out of closed-shell molecules are however not representable by a single determinant. Mathematically, singly excited singlet ($|\Psi_S\rangle$) and triplet ($|\Psi_T\rangle$) states corresponding to an excitation from orbital i to a from a closed-shell determinant $|\Phi\rangle$ are:

$$|\Psi_S\rangle = \frac{1}{\sqrt{2}}\left(a_a^\dagger a_i + a_{\bar{a}}^\dagger a_{\bar{i}}\right)|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|\Phi_i^a\rangle + |\Phi_{\bar{i}}^{\bar{a}}\rangle\right) \tag{9.9}$$

$$|\Psi_T\rangle = \frac{1}{\sqrt{2}}\left(a_a^\dagger a_i - a_{\bar{a}}^\dagger a_{\bar{i}}\right)|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|\Phi_i^a\rangle - |\Phi_{\bar{i}}^{\bar{a}}\rangle\right) \tag{9.10}$$

within the $M_s = 0$ manifold ($a^\dagger/a$ are second quantization creation/annihilation operators [14]). Equal contributions from both $|\Phi_i^a\rangle$ and $\left|\Phi_{\bar{i}}^{\bar{a}}\right\rangle$ is a consequence of both the up and down spins being equally likely to be excited. The broken symmetry $|\Phi_i^a\rangle$ determinant (when formed from spin-restricted orbitals) is thus an equal mixture of $|\Psi_S\rangle$ and $|\Psi_T\rangle$, indicating that ΔSCF is incapable of directly yielding spin-pure results. However, the triplet energy E_T can be accessed from the $M_s = \pm 1$ manifold, as the resulting configuration is well represented by a single determinant ($a_a^\dagger a_{\bar{i}}|\Phi\rangle = \left|\Phi_{\bar{i}}^a\right\rangle$ for the $M_s = 1$ case). The singlet energy E_S can then be obtained from approximate spin-projection [357] (AP):

$$E_S = 2E_{mixed} - E_T \tag{9.11}$$

This corresponds to $|\Phi_i^a\rangle$ being precisely halfway between $|\Psi_S\rangle$ and $|\Psi_T\rangle$ in energy. In general, $E_S > E_{mixed} > E_T$ is the usual ordering due to lack of an exchange stabilization term in the mixed configuration (vs the triplet).

In practice, E_{mixed} and E_T are usually separately optimized, resulting in two different sets of orbitals. Eqn 9.11 is subsequently carried out, amounting to projection after variation. This approach has been long used in the literature [264, 337, 358–360], and we refer to it as approximately projected ΔSCF (AP-ΔSCF). It is also possible to directly optimize Eqn. 9.11 for a single set of RO orbitals [361–363]. This has been described as restricted open-shell Kohn-Sham (ROKS) in the literature [361, 363], although it is distinct from normal RO KS-DFT calculations of high spin states of open-shell systems (where all unpaired spins point the same way [364]) or a related method proposed in Ref 365. The HF variant of ROKS has a long history of use [334–336] (often being described as "Open-Shell SCF" in older literature), as it represents a minimal, single configuration state function (CSF) approximation to open-shell singlets within multiconfigurational SCF (MCSCF) theory.

Both AP-ΔSCF and ROKS are however only applicable to singlet states with one broken electron pair. The utility of AP is limited for cases with a larger number of unpaired electrons, as each mixed determinant is subsequently a combination of more than two states [264], with several having the same $\langle S^2\rangle$. For instance, determinants with four unpaired electrons (two up and two down spins) are some mixture of one quartet, three triplets and two singlets. More general approaches are therefore required to obtain spin-pure results for such states. Such a procedure is described in Chapter 13 for the case of three electrons, and the protocol for an arbitrary number of electrons is laid out in Ref 366.

9.2.3 Formal Justification

It is often argued that a great advantage of TDDFT over OO-DFT is that the former is formally exact from the Runge-Gross theorem [250]. In contrast, there is no HK theorem for excited states [367], which indicates that there is no one-to-one mapping between the excited state density and the external potential (i.e. multiple external potentials can lead to the same excited state density). However, Ref 368 shows that there exists a one-to-one mapping between the kth excited state density and the external potential, as long as the external potential is purely Coulombic in character. It is therefore possible to define *subuniversal* functionals F_k that can predict the energy of the kth excited state for applied Coulombic potentials. More generally, it has also been shown [117] that every extremal density $\rho(\vec{r})$ of the exact ground state energy functional $E[\rho(\vec{r})]$ corresponds to an exact energy eigenstate (although the converse is not true, and not all excited states have densities that are stationary points of $E[\rho(\vec{r})]$). In fact, if $\rho_i(\vec{r})$ is an extremum of E, $E[\rho_i(\vec{r})]$ corresponds to the energy of the lowest energy state with density $\rho_i(\vec{r})$. This does indicate that use of OO-DFT approaches to extremize approximate ground state functionals could be effective in targeting excited states, as long as there is no state with the same density that is of lower energy. For instance, the lowest energy core-excited state should be well suited for OO-DFT, as there should not be any lower energy state that has similar density (due to the presence of the core-hole). In addition, Ref 368 shows that no two stationary states can have the same density *if* the applied external potential is Coulombic, which is the case for chemical systems not subjected to additional fields. Constrained search approaches based on minimizing parameters like the squared deviation of the energy from a target value, have also been proposed to formally justify excited state specific DFT for nondegenerate molecular bound states [369].

In practice therefore, states with a well-defined, unique electronic configuration (which could aid in having an unique density even when subjected to a non-Coulombic potential), that are "low energy" (likely characterizable by features such as a core-hole or subsystem charge) would be best approximated by OO-DFT. An alternative perspective would be to view E_{xc} as merely a tool for adding dynamic correlation to a truncated wavefunction.

Therefore, any state which can be reasonably approximated by one/few determinants that can be extremized via wavefunction theory, should be fairly well suited to OO-DFT as a qualitatively similar extremum is likely to be present and E_{xc} only contributes a minor (but often chemically essential) correction to the energy. In fact, this can be viewed as a reason why ground state KS-DFT has been immensely successful for single-reference species (where a single determinant from HF is a good approximation). As a corollary, excited states for which a reasonable truncated wavefunction model cannot be developed (or optimized) are likely to be challenging for OO-DFT approaches based upon existing KS functionals.

9.3 Core-level Spectroscopy

Spectroscopy of core electrons is an increasingly popular tool for studying chemical systems. Unlike valence levels, core electrons generally do not play a direct role in chemical bonding and remain localized around their respective nuclei. Furthermore, specific inner shells of a given element tend to have characteristic energies distinct from both other levels of the same element as well as all core-levels of other elements. Probing such electrons with X-ray or extreme ultraviolet (XUV) radiation is a convenient way to investigate the local chemical environment. Core-level spectroscopy has therefore been widely used to study both static properties [370–372] and chemical dynamics [373–375].

Theoretical modeling of core excited states is however a challenging task, as traditional quantum chemistry methods are typically geared towards understanding behavior of valence electrons. Excitations out of inner shells lead to substantial reorganization of the total electron density (the remaining core electron can have a more compact orbital due to reduced repulsion, not to mention the relaxation of the valence density in response), making them challenging for LR methods. TDDFT errors can readily span from underestimation by ∼10 eV [328] to overestimation by ∼10 eV (in the pure HF limit [376]), necessitating empirical translation of spectra for agreement with experiment (and motivating the development of specialized functionals optimized solely for core-level excitations [377]). Even equation of motion (EOM-) CCSD spectra often need to be shifted by a smaller amount (∼ 1-2 eV) to align experimental peaks with computed ones [378]. OO methods however can describe the relaxation of the density to all orders and therefore appear to be well suited for the task of modeling core-level spectra. This will be explored in Chapters 12 and 13.

9.4 Objectives

In this half of the dissertation, we will investigate the ability of DFT approaches to model excited states. We aim to both determine where existing approaches are insufficient, and develop methods that can handle problems considered to be challenging. In the process, we will cover a broad range of behaviors, spanning from excited states of stretched H_2 to core-level spectroscopy of radicals.

Chapter 10 studies the behavior of TDDFT and TDA type methods for describing bond dissociations in the excited state. The behavior of both spin-conserving excitations, as well as spin-flip transitions are examined. It is found that excited potential energy surfaces suffer from significant challenges at the Coulson-Fischer point (often having derivative discontinuities), with the lowest energy triplet state being described the most poorly. The results in this chapter indicate that these LR based approaches might not be effective at modeling reactive dynamics in the excited state.

In Chapter 11, we present a stable route, free from variational collapse, for excited state orbital optimization that focuses on the square of the gradient of the energy E vs orbital degrees of freedom $\vec{\theta}$. By defining the positive semidefinite quantity $\Delta = \left|\nabla_{\vec{\theta}} E\right|^2$, we can see that all stationary points of E correspond to global minima of Δ (and vice-versa). Minimization of Δ from an initial guess configuration should thus lead to the closest stationary point, as long as the gradient descent steps are sufficiently small. We show that this square gradient minimization (SGM) protocol only costs 2-3 times as much as ground state KS-DFT (per iteration), having similar computational cost as TDDFT. Furthermore, use of a general Lagrangian $\mathcal{L}$ instead of E would readily permit application of SGM to *any* quantum chemistry method like MP2 or CC doubles (CCD), and not just DFT.

SGM is subsequently employed to study the efficacy of OO-DFT in modeling core-level spectra. Chapter 12 presents the first use of ROKS for studying core excited states, which reveals that the modern SCAN functional has ~ 0.2 eV root mean square (RMS) error vs experimental excitation energies out of 1s orbitals of the elements C,N,O, and F. This is essentially semiquantitative performance, as the experimental energies have a typical uncertainty of ~ 0.1 eV. ROKS/SCAN is also able to predict spectra with good agreement with experiment, indicating that it is perhaps the most computationally efficient way to obtain core-level spectra that can be directly compared to experiment without need for empirical alignment. Good performance is also obtained for the 2p spectra of Si, P, S and Cl.

Examining the performance for core-level spectra of open-shell systems is harder, as excited states therein can have three or more unpaired electrons and therefore cannot be modeled with ROKS. We describe a general, CI inspired scheme to recouple an arbitrary number of unpaired spins from only single determinant KS energies in Chapter 13, which is further expanded upon in Ref 366. This recoupling approach is used in conjunction with SGM to study core-level spectra of radicals with OO-DFT. The results indicate the combination of OO and spin-purity in the recoupled approach permits it to be very effective in modeling core-excited states of radicals, including problems where EOM-CCSD qualitatively fails.

In summary, this half of the dissertation characterizes an area of TDDFT failure, presents a new OO-DFT orbital optimization algorithm, a recoupling scheme to describe multiconfigurational excitations, and applies OO-DFT to study core-level spectroscopy. It therefore aims to warn against the use of TDDFT for broad classes of problems, and argue that OO-DFT is able to resolve many of these issues for the same computational scaling.

Chapter 10

TDDFT beyond the Coulson-Fischer Point

10.1 Introduction

Linear-response time-dependent density functional theory [250, 251, 306, 379] (LR-TDDFT) is the most widely used excited state technique at present. The popularity of LR-TDDFT (henceforth simply referred to as TDDFT) is entirely a consequence of its computational affordability ($O(N^{2-3})$ cost versus molecule size [251]), which permits application to very large systems of hundreds of atoms. [380] Such species are well beyond the reach of more accurate wavefunction theory approaches like equation of motion coupled cluster [381, 382], or complete active space self-consistent field [383] (CASSCF) combined with corrections that include dynamic correlation [384, 385]. At the same time, TDDFT is considerably more accurate than the corresponding Hartree-Fock (HF) based wavefunction methods: specifically single excitation CI (CIS) [320] and time-dependent HF (TDHF) [37], which neglect dynamic correlation entirely.

In practice, TDDFT is plagued with many potential sources of error, despite having the potential to be formally exact [250] like ground state DFT [41]. TDDFT errors can roughly be viewed to originate from two sources: failure of the widely used adiabatic local density approximation [251, 379] (ALDA) and errors in the ground state DFT functional. The former generates large errors whenever the targeted state has large doubles (or higher order) character [326, 327], but is not expected to be a major problem for (almost) purely single excitations [251]. The latter remains a challenge despite the great accuracy of modern ground state density functionals [55, 80, 111, 226], as TDDFT tends to dramatically enhance relatively small ground state failures. The resulting excited state predictions are therefore considerably less reliable than the corresponding ground state calculations.

The most well known TDDFT failure is the systematic underestimation of excitation energies for charge-transfer (CT) and Rydberg states, on account of delocalization error [64, 83, 95] in the underlying functional [251, 270]. While delocalization error is an issue for ground state DFT as well, local exchange-correlation functionals predict a particularly poor description of long ranged particle-hole interaction [251, 270] in the linear-response limit. This becomes an issue whenever an electron undergoes a large spatial shift on account of the excitation, as is the case for CT or Rydberg states. The systematic underestimation (often on the order of 1-2 eV) can be mitigated (and sometimes over-corrected) via use of range separated hybrid functionals with considerable amounts of nonlocal exchange [93, 265, 386–388].

Another well-known failure of TDDFT stems from instability of the ground state Kohn-Sham DFT (KS-DFT) [48] solutions against mixing of occupied and virtual orbitals, which is often induced by static correlation. Similarities between the matrix diagonalized to obtain TDDFT excitation energies and the Hessian of the electronic energy against occupied-virtual mixing [15, 277, 284, 389] ensures that a negative eigenvalue in the latter (indicating an unstable ground state solution) often leads to a negative or imaginary “excitation" energy prediction by the former [251, 390]. A stable KS-DFT solution should not lead to such behavior, but sometimes ground state stability can only be achieved via artificial breakdown of spatial or spin symmetry.

A classic example is the breakdown of spin symmetry in unrestricted KS (UKS) and UHF calculations on closed-shell species, when single bonds are stretched beyond a point called the Coulson-Fischer (CF) point [212]. Spin symmetry breaks on account of the lowest triplet (T_1) state mixing with the singlet ground (S_0) state, and is consequently described as a ‘triplet instability’ [391]. The resulting spin-polarization leads to lower energies overall, but the corresponding asymmetric spin density cannot correspond to a spin-pure wavefunction. This form of spin contamination in UKS is irrelevant in the dissociation limit as the S_0 and T_1 states are then degenerate. However, it has consequences at shorter separations around the CF point where the fragment spins on the termini of the stretched bond are still partially coupled to each other. Qualitative success of unrestricted Hartree-Fock (UHF) in this regime [14] nonetheless suggests that UKS methods could yield smooth, qualitatively acceptable ground state potential energy surfaces (PES) for single bond dissociation, though some alarming failures by widely used functionals have been reported recently [218]. UKS calculations also guarantee size consistency in the ground state as the system energy at the single bond dissociation limit is identical to the sum of energies of isolated fragments.

It is possible to avoid spin contamination via using only spin-restricted (RKS) orbitals, but this would result in an unstable ground state solution that has an artificially elevated energy relative to the correct dissociation limit of isolated fragments (i.e. on account of artificial contamination from ionic dissociation products). Furthermore, TDDFT calculations on unstable RKS solutions would result in negative or imaginary triplet excitation energies [251, 392]. The lowest TDDFT singlet excitation energy is also known to spuriously go to zero at the dissociation limit for symmetric bonds when RKS orbitals are employed [393]. Overall, the benefits of spin-polarization for the ground state are well-recognized, and the associated limitations are understood as a usually acceptable price for smoothly joining accurate solutions at equilibrium (restricted) and dissociation (unrestricted spin-polarized). In wavefunction theory, use of multiple individually optimized HF determinants [341] as a basis for non-orthogonal CI (NOCI) [394] can go a long way towards restoring spin symmetry in both ground and excited states [395]. A similar approach is possible in DFT [396, 397], although the off-diagonal elements are not well-defined.

In this work, we explore the consequences of ground state spin-polarization in DFT (and HF) for excited states computed by TDDFT (and TDHF/CIS). We focus on single bond dissociations, beginning with the toy problem of H_2 in a minimum basis for HF and CIS/TDHF. We then move to more realistic basis sets in a variety of stretched single bond systems, including TDDFT as well. A variety of interesting artifacts are found in these results, which have their origins in the neglect of all double excitations, and in the link between characterizing orbital (in)stability and excited states. We suspect that some of these results have been seen by researchers before (some have been mentioned in Ref 398, for instance), but we believe there is no careful study examining this issue in detail. The results indicate that a great deal of caution is needed when using unrestricted orbitals for TDDFT/TDHF/CIS (as well as related methods like CIS(D) [399] or CC2 [400]) in closed shell systems beyond the CF point!

10.2 Spin-flipping Excitations within TDDFT/TDHF

A full description of the TDDFT equations is provided in 9.1. Eqn 9.7 offers a physical interpretation of the indices ia and jb as representations of excitations from occupied spin orbital ϕ_i to virtual spin orbital ϕ_a and from occupied spin orbital ϕ_j to virtual spin orbital ϕ_b, respectively. This interpretation can be generalized beyond CIS to TDHF and even to TDDFT, despite the fictitious nature of KS orbitals.

Labeling each occupied-virtual pair by their spins alone (i.e. $\alpha\alpha, \alpha\beta$ etc.) and integrating out spin degrees of freedom in Eqn 9.2, we find that the structure of the $\mathbf{A}$ matrix is :

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{\alpha\alpha,\alpha\alpha} & \mathbf{A}_{\alpha\alpha,\beta\beta} & 0 & 0 \\ \mathbf{A}_{\beta\beta,\alpha\alpha} & \mathbf{A}_{\beta\beta,\beta\beta} & 0 & 0 \\ 0 & 0 & \mathbf{A}_{\alpha\beta,\alpha\beta} & 0 \\ 0 & 0 & 0 & \mathbf{A}_{\beta\alpha,\beta\alpha} \end{pmatrix} \tag{10.1}$$

This leaves

$$\mathbf{A}_{M_S=0} = \begin{pmatrix} \mathbf{A}_{\alpha\alpha,\alpha\alpha} & \mathbf{A}_{\alpha\alpha,\beta\beta} \\ \mathbf{A}_{\beta\beta,\alpha\alpha} & \mathbf{A}_{\beta\beta,\beta\beta} \end{pmatrix} \tag{10.2}$$

as the spin-conserving, $M_S = 0$ block, since the spin of the occupied electron being excited to a virtual orbital does not change. On the other hand, $\mathbf{A}_{\alpha\beta,\alpha\beta}$ and $\mathbf{A}_{\beta\alpha,\beta\alpha}$ represent spin-flipping $M_S = \mp 1$ blocks, as they depict the transition from an α occupied to a β virtual and the reverse, respectively. Similarly, we find that:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{\alpha\alpha,\alpha\alpha} & \mathbf{B}_{\alpha\alpha,\beta\beta} & 0 & 0 \\ \mathbf{B}_{\beta\beta,\alpha\alpha} & \mathbf{B}_{\beta\beta,\beta\beta} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{B}_{\alpha\beta,\beta\alpha} \\ 0 & 0 & \mathbf{B}_{\beta\alpha,\alpha\beta} & 0 \end{pmatrix} \tag{10.3}$$

It can immediately be seen that the spin-conserving block

$$\mathbf{B}_{M_S=0} = \begin{pmatrix} \mathbf{B}_{\alpha\alpha,\alpha\alpha} & \mathbf{B}_{\alpha\alpha,\beta\beta} \\ \mathbf{B}_{\beta\beta,\alpha\alpha} & \mathbf{B}_{\beta\beta,\beta\beta} \end{pmatrix} \tag{10.4}$$

is independent of the spin-flipping block, like $\mathbf{A}_{M_S=0}$. The eigenvalues obtained from the spin-conserving block alone are a subset of the exact solutions to the full Eqn 9.6, and standard TDDFT/TDHF procedures typically focus only on these excitations. The spin-flipping block nonetheless does contain physical content, and is essential for obtaining states with different M_S than the reference. For instance, the $M_S = \pm 1$ triplets for a molecule with a singlet ground state can only be obtained from the spin-flipping block, while the $M_S = 0$ state can be obtained from the spin-conserving block. It is also worth noting that the spin-flipping blocks of the **A** and **B** matrices are involved in determining whether UHF solutions are stable against spin-flipping orbital rotations to Generalized HF (GHF) solutions [15].

Excluding the spin-conserving $M_S = 0$ block that perfectly separates from the rest, we have Eqn 9.1 reduce to:

$$\begin{pmatrix} \mathbf{A}_{\alpha\beta,\alpha\beta} & 0 & 0 & \mathbf{B}_{\alpha\beta,\beta\alpha} \\ 0 & \mathbf{A}_{\beta\alpha,\beta\alpha} & \mathbf{B}_{\beta\alpha,\alpha\beta} & 0 \\ 0 & \mathbf{B}_{\alpha\beta,\beta\alpha} & \mathbf{A}_{\alpha\beta,\alpha\beta} & 0 \\ \mathbf{B}_{\beta\alpha,\alpha\beta} & 0 & 0 & \mathbf{A}_{\beta\alpha,\beta\alpha} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\alpha\beta} \\ \mathbf{X}_{\beta\alpha} \\ \mathbf{Y}_{\alpha\beta} \\ \mathbf{Y}_{\beta\alpha} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & -\mathbf{1} & 0 \\ 0 & 0 & 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\alpha\beta} \\ \mathbf{X}_{\beta\alpha} \\ \mathbf{Y}_{\alpha\beta} \\ \mathbf{Y}_{\beta\alpha} \end{pmatrix} \tag{10.5}$$

This again can be separated into two independent blocks:

$$\begin{pmatrix} \mathbf{A}_{\beta\alpha,\beta\alpha} & \mathbf{B}_{\beta\alpha,\alpha\beta} \\ \mathbf{B}_{\alpha\beta,\beta\alpha} & \mathbf{A}_{\alpha\beta,\alpha\beta} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\beta\alpha} \\ \mathbf{Y}_{\alpha\beta} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\beta\alpha} \\ \mathbf{Y}_{\alpha\beta} \end{pmatrix} \tag{10.6}$$

$$\begin{pmatrix} \mathbf{A}_{\alpha\beta,\alpha\beta} & \mathbf{B}_{\alpha\beta,\beta\alpha} \\ \mathbf{B}_{\beta\alpha,\alpha\beta} & \mathbf{A}_{\beta\alpha,\beta\alpha} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\alpha\beta} \\ \mathbf{Y}_{\beta\alpha} \end{pmatrix} = \omega \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{\alpha\beta} \\ \mathbf{Y}_{\beta\alpha} \end{pmatrix} \tag{10.7}$$

The indices in Eqn 10.7 can be rearranged to yield:

$$\begin{pmatrix} \mathbf{A}_{\beta\alpha,\beta\alpha} & \mathbf{B}_{\beta\alpha,\alpha\beta} \\ \mathbf{B}_{\alpha\beta,\beta\alpha} & \mathbf{A}_{\alpha\beta,\alpha\beta} \end{pmatrix} \begin{pmatrix} \mathbf{Y}_{\beta\alpha} \\ \mathbf{X}_{\alpha\beta} \end{pmatrix} = \omega \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{Y}_{\beta\alpha} \\ \mathbf{X}_{\alpha\beta} \end{pmatrix} \tag{10.8}$$

which is *nearly* identical in structure to Eqn 10.6, save a sign. It is therefore evident that Eqns 10.6 and 10.7 share the same eigenvectors, and the corresponding eigenvalues differ only by a sign. Although only positive eigenvalues (i.e. excitation energies) have physical meaning, the negative eigenvalues of Eqn 10.6 yield the positive eigenvalues of Eqn 10.7, and so it suffices to fully solve Eqn 10.6 alone to have all the excitation energies from the spin flipping block. Physically, this can be interpreted as energies for all de-excitations with $M_S = 1$ being the negative of excitation energies with $M_S = -1$ (and vice versa). TDA/CIS here is trivially achieved by diagonalizing the independent $M_S = \pm 1$ blocks $\mathbf{A}_{\beta\alpha,\beta\alpha}$ and $\mathbf{A}_{\alpha\beta,\alpha\beta}$ blocks separately.

An important difference between the spin-flipped blocks arising from HF and typical KS solutions also merits a mention. To linear-response, spin-flipping excitations only affect the off-diagonal $\mathbf{P}_{\alpha\beta}/\mathbf{P}_{\beta\alpha}$ blocks of the one particle density matrix $\mathbf{P}$, and consequently do not affect the electron density at all. Collinear exchange-correlation kernels f_{xc} will therefore have zero contribution from the local component of the exchange-correlation functional. In other words, matrix elements involving f_{xc} will be zero in the spin-flipping block for purely local functionals like PBE [57], and only contributions from HF exchange will count for hybrid functionals like PBE0 [91] or LRC-ωPBEh [179]. This means that the local exchange-correlation contribution to Eqn. 9.2 via orbital energy differences will go uncorrected in KS theory, which (as we will demonstrate) leads to unusual behavior for $M_S = \pm 1$ solutions for TDDFT/TDA relative to CIS. Furthermore, there will be no contributions from local exchange-correlation terms to the $\mathbf{B}$ matrix within the spin-flip block, rendering TDA identical to full TDDFT for local functionals.

10.3 CIS and TDHF for Stretched H_2 in a Minimal Basis.

Let us first consider the instructive toy model of minimal basis (STO-3G) [401] H_2, which contains 2 spatial orbitals and 2 electrons. Six determinants are consequently possible, and are illustrated in Fig 10.1. The $M_S = 0$ subspace has four determinants and the $M_S = \pm 1$ subspaces have one each. These determinants can be formed from either restricted (R) or unrestricted spin-polarized (U) orbitals. The R orbitals are bonding ($|\sigma\rangle$) and antibonding ($|\sigma^*\rangle$) respectively, while the U orbitals at dissociation are atomic orbitals ($|1\mathrm{s}_\mathrm{A}\rangle$ and $|1\mathrm{s}_\mathrm{B}\rangle$). Exact full configuration interaction (FCI) is invariant to the choice of orbitals, but HF and TDHF/CIS show critical differences between the R and U cases.

FCI in the rank-4 $M_S = 0$ manifold yields three singlets and a triplet at all possible distances. At equilibrium, these states are (roughly) a $|\sigma\bar{\sigma}\rangle$ singlet ground state ($\mathrm{X}^1\Sigma_\mathrm{g}$), (roughly) a $|\sigma^*\bar{\sigma^*}\rangle$ doubly excited singlet ($\mathrm{A}^1\Sigma_\mathrm{g}$), along with a singlet ($^1\Sigma_\mathrm{u}$) and a triplet ($^3\Sigma_\mathrm{u}$) resulting from linear combinations of the single excitations $|\sigma\bar{\sigma^*}\rangle$ and $|\sigma^*\bar{\sigma}\rangle$. The two additional states with $M_S = \pm 1$ complete the triplet manifold. In the dissociation limit, the lowest $^1\Sigma_\mathrm{g}$ state and $^3\Sigma_\mathrm{u}$ become degenerate, arising from the four possible ways in which the spins on two isolated H atoms can couple. There are also two degenerate ($^1\Sigma_\mathrm{g}$ and $^1\Sigma_\mathrm{u}$) higher energy CT states corresponding to superpositions of $^-\mathrm{H}\cdots\mathrm{H}^+$ and $^+\mathrm{H}\cdots\mathrm{H}^-$.

It is instructive to consider the behavior of HF/CIS against this exact behavior. At

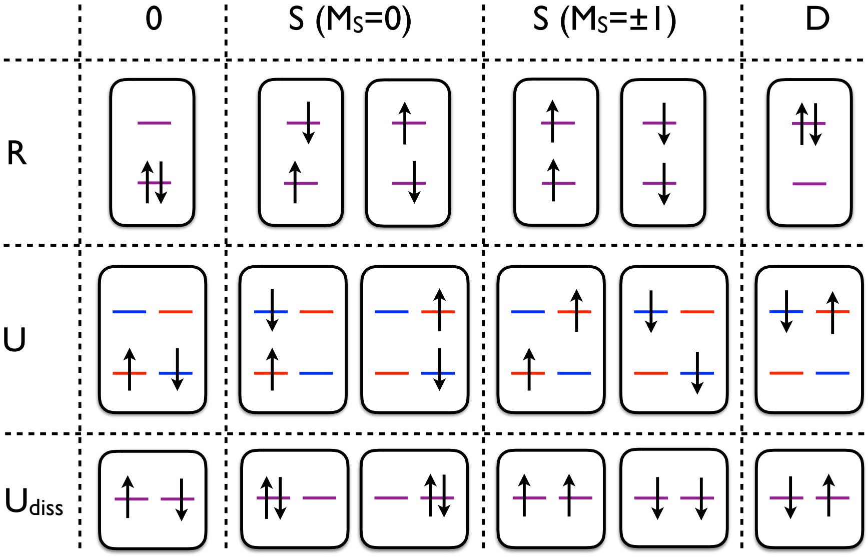
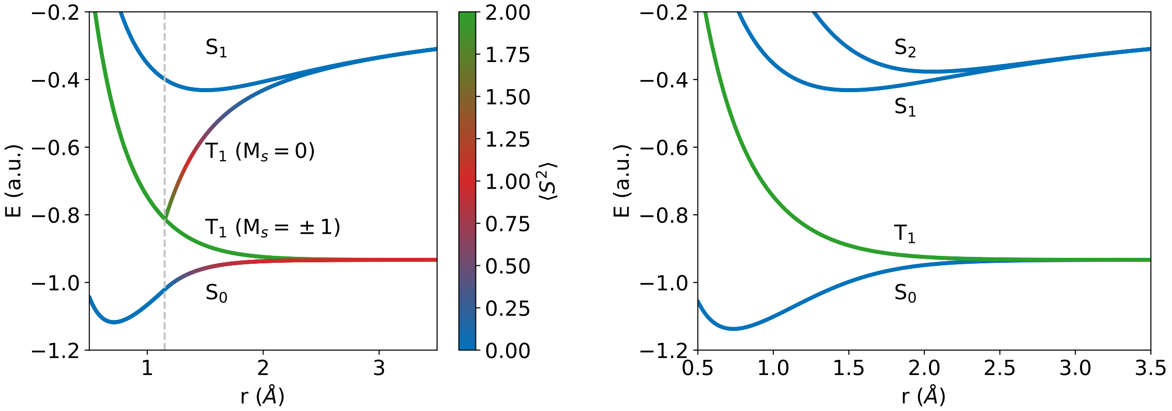


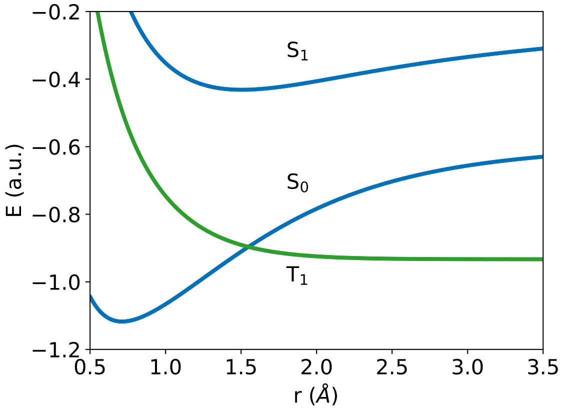
Figure 10.1: All possible Slater determinants for minimal basis H_2, using restricted (R), spin-polarized unrestricted (U), and dissociation-limit spin-polarized unrestricted (U_{diss}) orbital representations for the ground state determinant (0). Aside from the ground state, the four possible single (S) excitations, and the one double (D) excitation are also depicted.

internuclear distances smaller than the CF point, the stable HF ground state is $|\sigma\bar{\sigma}\rangle$ (as seen in Fig 10.1) and has no spin-polarization. There are 2 possible $M_S = 0$ single substitutions ($|\sigma\bar{\sigma^*}\rangle$ and $|\sigma^*\bar{\sigma}\rangle$, as shown in Fig 10.1), which arise from $\sigma \to \sigma^*$ transitions. Diagonalizing the CIS Hamiltonian therefore gives a singlet and a triplet state (which are positive and negative linear combinations of $|\sigma\bar{\sigma^*}\rangle$ and $|\sigma^*\bar{\sigma}\rangle$ with equal weights). The $M_S = \pm 1$ states (which are degenerate with the $M_S = 0$ triplet) complete the triplet manifold. The situation corresponds very well to FCI (indeed, the ${}^1\Sigma_u$ and ${}^3\Sigma_u$ levels are exact), save the absence of the doubly excited singlet, which is beyond the scope of a 'singles-only' method like CIS.

Let us now consider the (unrestricted) dissociation limit. The stable UHF ground state has an electron localized on each atom (lowest panel of Fig 10.1), and is spin-contaminated (equal parts singlet and triplet). This spin-polarized UHF state is energetically preferred over RHF due to the absence of spurious ionic (CT) contributions [14] in the former, which make up 50% of the wavefunction in the latter. There are two $M_S = 0$ single substitutions (Fig 10.1) which are singlet with CT character (i.e. $H^- \cdots H^+$ and $H^+ \cdots H^-$). These two singly excited CT determinants are noninteracting, making them UCIS eigenstates which exactly match the FCI CT states. The remaining singles are the covalent $M_S = \pm 1$ states, which also match FCI. So, at dissociation, UCIS is exact for the $M_S = \pm 1$ components of the FCI triplet, but entirely fails to describe the $M_S = 0$ sub-level of the state! A portion of the $M_S = 0$ triplet survives via mixing with the ground state to induce spin polarization, while the remainder lies in the omitted doubly excited D determinant.



(a) UHF and UCIS (b) Exact



(c) RHF and RCIS

Figure 10.2: HF and CIS PESs of H_2 in the minimal basis, compared to FCI. The $M_S = 0$ UCIS T_1 state has a pronounced kink at the CF point (marked by the pale gray dashed line) and subsequently ascends to the CT dissociation limit. However, the $M_S = \pm 1$ T_1 states continue on to the proper dissociation limit of neutral atoms in the ground state. The RCIS T_1 and S_1 states are exact, but the S_0 state has significant CT character, resulting in a spurious elevation of energy.

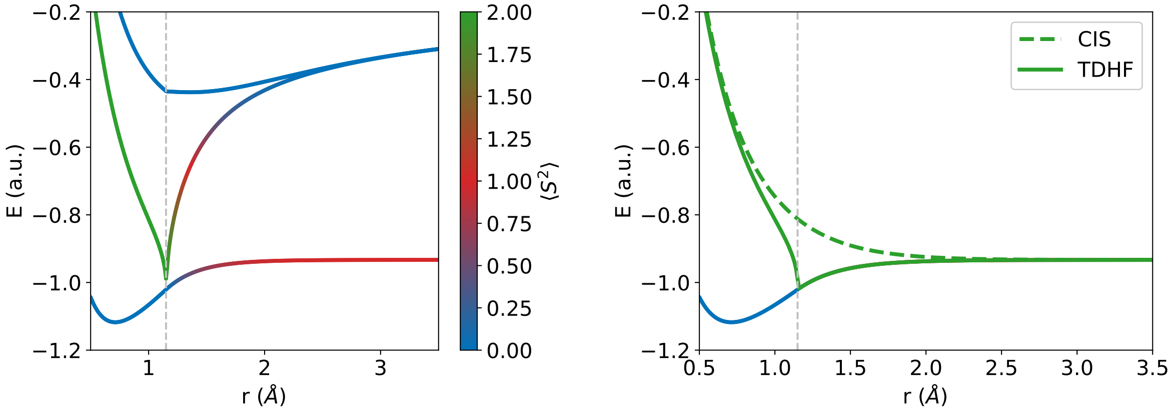
The complete FCI PESs are shown in Fig 10.2b, which can be compared against the UHF/UCIS results in Fig 10.2a and RHF/RCIS results in Fig 10.2c. The FCI T_1 surface is nonbonding, and the energy decreases monotonically with internuclear separation r. As already discussed above, the RCIS T_1 is exact. However, spin polarization in the UHF ground state leads to a very sharp and unphysical kink in the UCIS $M_S = 0$ T_1 state at the CF point (spuriously suggesting a local minimum), followed by a monotonic rise in energy to the CT dissociation limit, versus the desired neutral atom limit. This state therefore changes

	RHF	UHF
Spatial molecular orbitals		
Occupied	$\lvert\sigma\rangle = \dfrac{\lvert 1\mathrm{s_A}\rangle + \lvert 1\mathrm{s_B}\rangle}{\sqrt{2(1+\langle 1\mathrm{s_A}\vert 1\mathrm{s_B}\rangle)}}$	$\lvert\alpha_1\rangle = \cos\theta\,\lvert\sigma\rangle + \sin\theta\,\lvert\sigma^*\rangle$ $\lvert\beta_1\rangle = \cos\theta\,\lvert\sigma\rangle - \sin\theta\,\lvert\sigma^*\rangle$
Virtual	$\lvert\sigma^*\rangle = \dfrac{\lvert 1\mathrm{s_A}\rangle - \lvert 1\mathrm{s_B}\rangle}{\sqrt{2(1-\langle 1\mathrm{s_A}\vert 1\mathrm{s_B}\rangle)}}$	$\lvert\alpha_2\rangle = \cos\theta\,\lvert\sigma^*\rangle - \sin\theta\,\lvert\sigma\rangle$ $\lvert\beta_2\rangle = \cos\theta\,\lvert\sigma^*\rangle + \sin\theta\,\lvert\sigma\rangle$
HF/CIS electronic states		
$\lvert\mathrm{S_0}\rangle$	$\lvert\sigma\bar{\sigma}\rangle$	$\lvert\alpha_1\bar{\beta}_1\rangle = \cos^2\theta\,\lvert\mathrm{S_0}\rangle_{\mathrm{RHF}} - \sin^2\theta\,\lvert\mathrm{S_2}\rangle_{\mathrm{RCIS}} - \dfrac{1}{\sqrt{2}}\sin 2\theta\,\lvert\mathrm{T_1}\rangle_{\mathrm{RCIS}}^{\mathrm{M_S=0}}$
$\lvert\mathrm{T_1}\rangle^{\mathrm{M_S=1}}$	$\lvert\sigma\sigma^*\rangle$	$\lvert\alpha_1\alpha_2\rangle = \lvert\mathrm{T_1}\rangle_{\mathrm{RCIS}}^{\mathrm{M_S=1}}$
$\lvert\mathrm{T_1}\rangle^{\mathrm{M_S=0}}$	$\dfrac{\lvert\sigma\bar{\sigma^*}\rangle - \lvert\sigma^*\bar{\sigma}\rangle}{\sqrt{2}}$	$\dfrac{\lvert\alpha_1\bar{\beta}_2\rangle - \lvert\alpha_2\bar{\beta}_1\rangle}{\sqrt{2}} = \cos 2\theta\,\lvert\mathrm{T_1}\rangle_{\mathrm{RCIS}}^{\mathrm{M_S=0}} + \dfrac{\sin 2\theta}{\sqrt{2}}\left(\lvert\mathrm{S_0}\rangle_{\mathrm{RHF}} + \lvert\mathrm{S_2}\rangle_{\mathrm{RCIS}}\right)$
$\lvert\mathrm{T_1}\rangle^{\mathrm{M_S=-1}}$	$\lvert\bar{\sigma}\bar{\sigma^*}\rangle$	$\lvert\bar{\beta}_1\bar{\beta}_2\rangle = \lvert\mathrm{T_1}\rangle_{\mathrm{RCIS}}^{\mathrm{M_S=-1}}$
$\lvert\mathrm{S_1}\rangle$	$\dfrac{\lvert\sigma\bar{\sigma^*}\rangle + \lvert\sigma^*\bar{\sigma}\rangle}{\sqrt{2}}$	$\dfrac{\lvert\alpha_1\bar{\beta}_2\rangle + \lvert\alpha_2\bar{\beta}_1\rangle}{\sqrt{2}} = \lvert\mathrm{S_1}\rangle_{\mathrm{RCIS}}$
$\lvert\mathrm{S_2}\rangle$	$\lvert\sigma^*\bar{\sigma^*}\rangle$	$\lvert\alpha_2\bar{\beta}_2\rangle = \cos^2\theta\,\lvert\mathrm{S_2}\rangle_{\mathrm{RHF}} - \sin^2\theta\,\lvert\mathrm{S_0}\rangle_{\mathrm{RCIS}} - \dfrac{1}{\sqrt{2}}\sin 2\theta\,\lvert\mathrm{T_1}\rangle_{\mathrm{RCIS}}^{\mathrm{M_S=0}}$

Table 10.1: Analytic representation of molecular orbitals and electronic states predicted by HF/CIS for minimal basis H_2. $\lvert 1\mathrm{s_A}\rangle$ and $\lvert 1\mathrm{s_B}\rangle$ are the atomic basis functions. Note that the doubly excited $\lvert\mathrm{S_2}\rangle$ state is not directly predicted by CIS, but is easily identifiable by virtue of being the only state in the Hilbert space not directly predicted by HF/CIS.

character from triplet at the CF point to CT singlet as dissociation approaches, consistent with the analysis given above. In contrast, the UCIS $M_S = \pm 1$ components of the $\mathrm{T_1}$ state remain exact.

Additional insight into the curves can be gained by monitoring spin-polarization in the stable UHF determinant via a parameter $\theta \in \left[0, \frac{\pi}{4}\right]$ for orbital mixing. Table 10.1 provides the analytic forms for the orbitals and the HF/CIS states in terms of θ. $\theta = 0$ for $r < r_{\mathrm{CF}}$ (RHF regime), and $\theta = \frac{\pi}{4}$ at the dissociation limit (as described in Sec 3.8.7 of Ref 14). The $\mathrm{T_1}$ state mixes with the singlet ground state beyond the CF point (as can be seen from Table 10.1), resulting in a spin-contaminated state. Consequently, ground state $\langle S^2\rangle = \sin^2 2\theta$, which shows monotonic change from a pure singlet ($\langle S^2\rangle = 0$ for $\theta = 0$) at $r = r_{\mathrm{CF}}$ to equal singlet-triplet mixture ($\langle S^2\rangle = 1$ for $\theta = \frac{\pi}{4}$) at the dissociation limit. Similarly, the analytic continuation of the $M_S = 0$ UCIS $\mathrm{T_1}$ state has $\langle S^2\rangle = 2\cos^2 2\theta$. It is therefore a pure triplet in the RHF regime ($\langle S^2\rangle = 2$ for $\theta = 0$), spin polarizes past the CF point to become a singlet-triplet mixture, and ultimately becomes a pure singlet at dissociation ($\langle S^2\rangle = 0$ for $\theta = \frac{\pi}{4}$), showing a complete change of character. In contrast, the spin-polarized continuation of the $M_S = \pm 1$ components of the RCIS triplet remain exact (and smooth) beyond the CF point because they are already exact as single determinants. Likewise, Table 10.1 shows the lowest singlet excited state (S_1) surface is exact, smoothly changing from a valence to CT excited state without any triplet contamination.



(a) $M_S = 0$ block

(b) Spin-flipping block

Figure 10.3: TDHF excited states for minimal basis H_2 ($\langle S^2\rangle$ values correspond to the corresponding CIS states). The $M_S = \pm 1$ CIS solution is also supplied for comparison.

It is also instructive to consider the behavior of TDHF proper to see how the inclusion of the **B** matrix affects CIS results. Fig 10.3a reveals that the TDHF states show essentially the same general picture as CIS within the spin-conserving $M_S = 0$ block. The kink in the T_1 state is even more pronounced, as the surface effectively funnels down to the ground state (i.e. zero excitation energy) at the CF point, before an even more steep ascent to the incorrect dissociation limit. This is a consequence of $\mathbf{A} + \mathbf{B}$ having a zero eigenvalue at the CF point due to onset of spin-polarization induced instability, which leads to a zero eigenvalue for Eqn. 9.6. The S_1 state also has a weak kink in the CF point (unlike the case of CIS, where it was exact), but both excited states go to the exact CT dissociation limit. Overall, the performance of full TDHF is somewhat worse than the already poor performance of CIS, consistent with earlier observations [316].

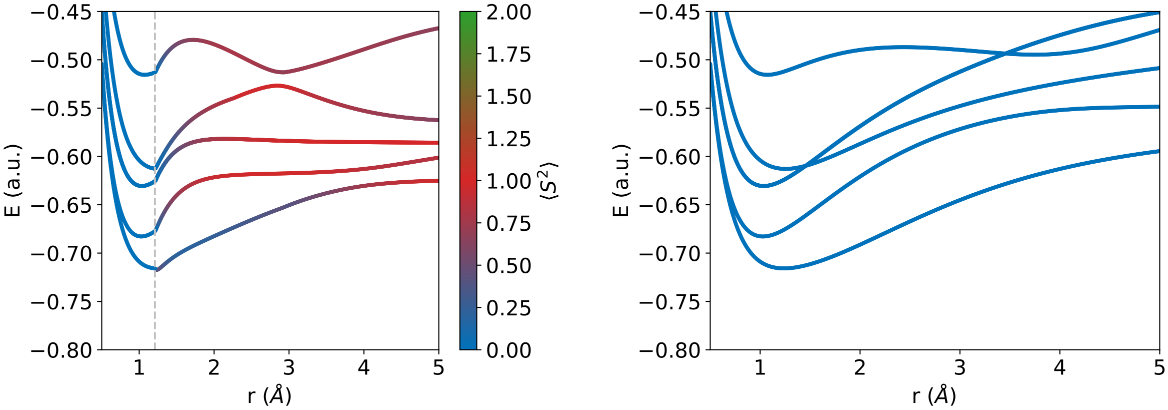
The behavior of TDHF in the spin-flip blocks is quite distinct, as made evident by Fig 10.3b. The spin-flip T_1 states are degenerate with the spin-conserving one prior to spin-polarization, going to the expected zero excitation energy at the CF point and exhibiting a derivative discontinuity therein. They however subsequently remain degenerate with the UHF ground state (i.e. have zero excitation energy). An analytic proof for the zero spin-flip TDHF excitation energy for this toy model is supplied in Sec 10.8. A more general argument however can be derived from GHF stability theory. The direction of the spin-density induced by the spin-polarization is arbitrary within GHF theory (unlike in UHF where it is constrained to be along the z direction), and therefore orbital rotations that break $\hat{S}_x$ and $\hat{S}_y$ symmetries do not have any associated energy barrier or restoring force [402]. Consequently, the GHF stability Hessian [15] has two zero eigenvalues corresponding to these orbital rotation normal modes. These lead to four zero eigenvalues in Eqn 10.5 [402] and two zero eigenvalues in Eqn 10.6. These 0 eigenvalues are natural continuations of the pre-CF point T_1 curves and consequently can be viewed as T_1 excitation energies even beyond the CF point .

Let us briefly summarize this minimal basis H_2 story of disaster and success beyond the CF point. The UHF ground S_0 state separates correctly with spin-polarization. However, spin-polarization causes the CIS T_1 to separate correctly only in the $M_S = \pm 1$ sub-levels, while the $M_S = 0$ sub-level becomes a CT singlet at separation. The CIS S_1 state nonetheless remains exact. TDHF further worsens the CIS results, with the $M_S = 0$ T_1 solution possessing a more pronounced kink that spuriously connects it to the S_0 state. A kink is also induced in the previously perfect S_1 state. TDHF within the spin-flip block reveals that while the T_1 states remain degenerate with the $M_S = 0$ solution until the CF point, they separate afterwards, with the continuations of the spin-flip T_1 states becoming degenerate with the UHF ground state post-spin polarization.

10.4 Stretched H_2 in a Larger Basis.

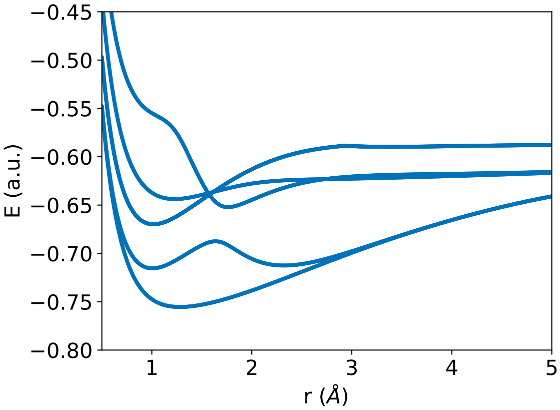
We next consider the behavior of the singlet and triplet excited states (and their continuations past the CF point) for the larger aug-cc-pVTZ basis, as shown in Figs. 10.4 and 10.5 respectively. Both singlet and triplet surfaces are significantly impacted by the CF point. Fig 10.4a shows that there is a sudden increase in energy for the singlet surfaces right beyond the CF point, with a clear first derivative discontinuity at $r = r_{\mathrm{CF}}$. All depicted singlets develop some spin contamination as well (though any purely CT state would not have this issue). A comparison to exact surfaces in Fig 10.4a seems to suggest that all the affected states switch over to entirely new asymptotic regimes relative to their initial trajectory prior to the CF point, often accompanied by a dramatic change in the PES curvature. Spin polarization therefore seems to connect two distinct surfaces beyond the CF point, via the kink.

Similar behavior is observed for the low lying $M_S = 0$ triplet surfaces (Fig 10.5a), where shallow minima are effectively deepened relative to the dissociation limit via spin polarization. The $M_S = 0$ T_1 state shows artifacts that resemble those seen in the minimal basis, slowly losing triplet character to become a CT state at intermediate separation. At long separations, it again spin polarizes to form an excited state localized on a single H atom with $\langle S^2 \rangle = 1$, which is lower in energy than the CT state. All character of the exact T_1 state (where both electrons are essentially on linear combinations of 1s orbitals) is consequently erased. In contrast, Fig 10.5b shows that the $M_S = \pm 1$ T_1 states remain qualitatively acceptable in larger basis sets and go to the correct dissociation limit. Higher energy triplets still have a kink at the CF point on account of spin polarization, and there is not much improvement in energy relative to $M_S = 0$ subspace (cf. Fig 10.5a), despite the states now being perfectly spin pure. The dramatic improvement in the quality of the T_1 state is nonetheless very promising, as it is often the principal actor in photodissociation.



(a) UCIS

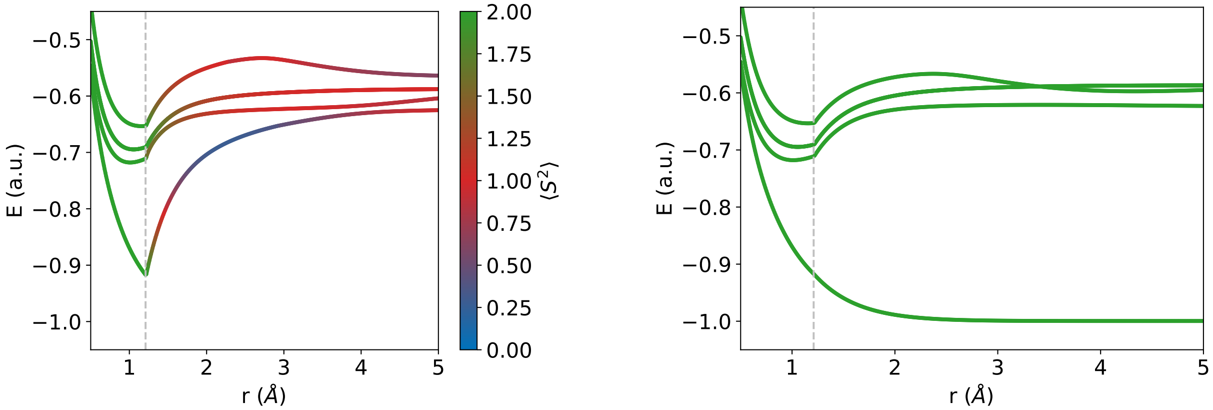
(b) RCIS



(c) Exact

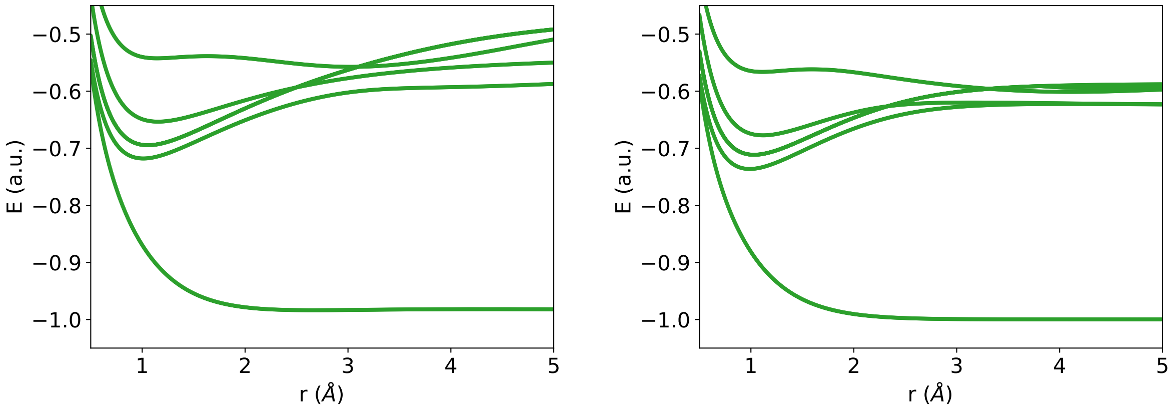
Figure 10.4: Low lying CIS singlet excited state PESs of H_2 in the aug-cc-pVTZ basis compared to exact results. Small state crossing induced discontinuities might be present on the top surface. The UCIS singlet states are spin-contaminated beyond the CF point, but their analytical continuation is still followed to the dissociation limit. The S_0 ground state has not been depicted, for clarity.

It is worth noting that the RCIS singlet and triplet surfaces (depicted in Figs. 10.4b and 10.5c respectively) are smooth, and appear to be mostly physical in comparison to the exact surfaces, despite being somewhat higher in energy (especially in the dissociation limit) due to missing correlation. It also appears that many excited states are slower to reach their asymptotic limits (relative to UCIS/FCI), as evidenced by relatively large slopes at even 5 Å separation. This could be a consequence of CT character of the RHF reference being carried over to the excited states, as CT state energies asymoptotically decay as r^{-1} (vs valence excitation energies, which decay exponentially to the asymptotic limit, like the fragment wavefunction overlap). This is however difficult to characterize for H_2 as there is no *net* charge transfer, and only a two electron property (like a pair distribution function) would therefore be able to reveal whether the excited RCIS states have spurious CT character like



(a) UCIS $M_S = 0$.

(b) UCIS $M_S = \pm 1$.



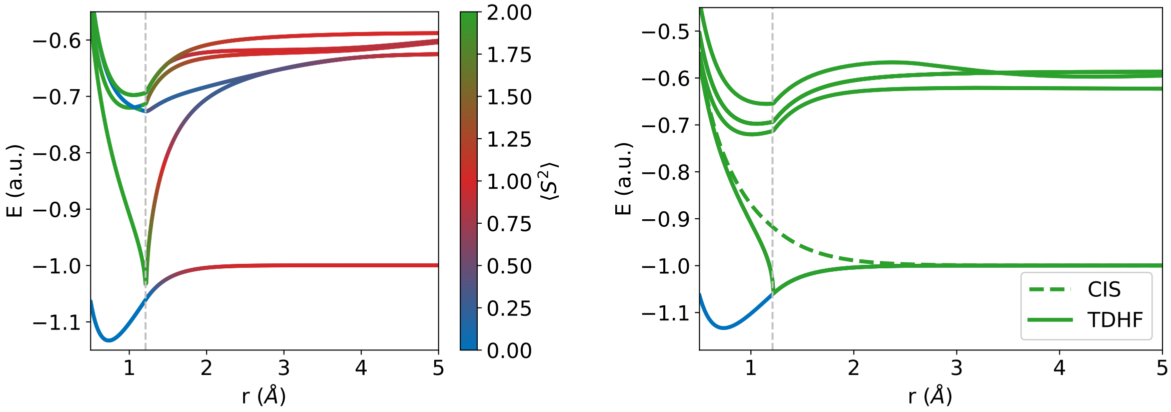
(c) RCIS

(d) Exact

Figure 10.5: The CIS triplet excited state PESs of H_2 in the aug-cc-pVTZ basis compared to exact results. Small state crossing induced discontinuities might be present on the top surface. The UCIS $M_S = 0$ triplet states are spin-contaminated beyond the CF point, but their analytical continuation is still followed to the dissociation limit.

the RHF reference. The T_1 state however asymptotes at a reasonable rate and appears to get close to the correct dissociation limit of independent ground state H atoms (though is too high in energy by 0.47 eV). The qualitatively acceptable performance of the RCIS T_1 excited state however comes at the cost of a severely compromised RHF S_0 ground state, where spurious CT contributions drive it above the T_1 state by 0.21 a.u. (5.7 eV) at the dissociation limit!

Moving on to TDHF, Fig 10.6a shows that the $M_S = 0$ TDHF energy surfaces look similar to the CIS ones shown earlier in Figs 10.4a and 10.5a. The quality of the T_1 surface is worse, as was the case for the minimal basis. Fig 10.6b shows that the $M_S = \pm 1$ T_1 TDHF solutions become degenerate with the UHF S_0 state past the CF point, as was the case with minimal basis as well. The $M_S = \pm 1$ higher excited states however are very similar to those obtained from CIS, including kinks at the CF point. Restricted orbitals offer no benefits for

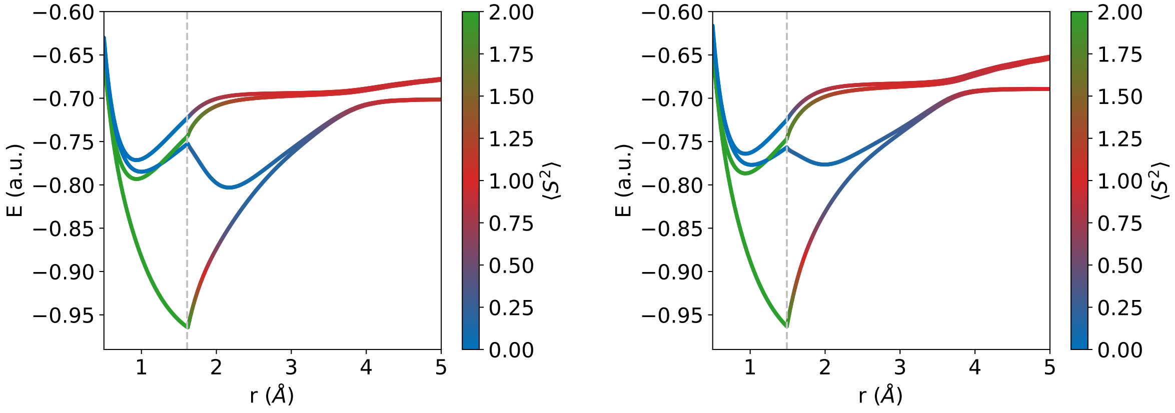


(a) $M_S = 0$ block.

(b) Spin-flipping block.

Figure 10.6: TDHF excited state PESs of H_2 in the aug-cc-pVTZ basis. The states have been labeled with the $\langle S^2 \rangle$ of the corresponding CIS state. The $M_S = \pm 1$ CIS state has also been supplied as a reference.

TDHF, as the instability in the RHF solution beyond the CF point causes the T_1 excitation energy to become imaginary (via Eqn 9.6). The dissociation limit S_1 excitation energy is also known to spuriously go to zero for H_2, and for all other symmetric bond dissociations [393]. Use of full TDHF (TDDFT) is therefore unlikely to lead to any qualitative improvements for excited state PESs around and beyond the CF point, relative to CIS (TDA).

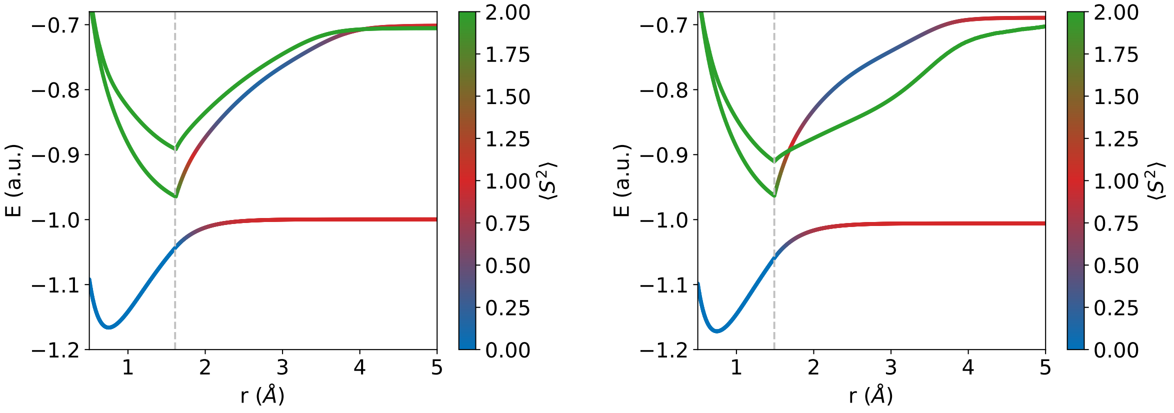


(a) PBE

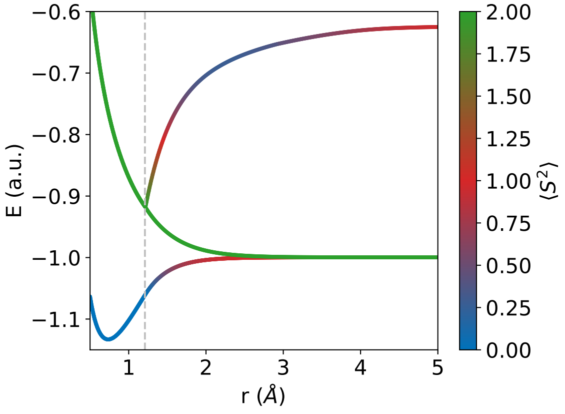
(b) LRC-ωPBEh

Figure 10.7: TDDFT/TDA (within the $M_S = 0$ subspace) for stretched H_2/ aug-cc-pVTZ.

The observations discussed so far stem from the general formalism of TDDFT versus any specific issues with the HF functional. It therefore seems that CIS/TDHF results should be transferable to other density functionals, with minor adjustments. We demonstrate this by providing TDDFT/TDA PES for the local PBE [57] and range separated hybrid LRC-ωPBEh [179] functionals (within the $M_S = 0$ subspace) in Fig 10.7. We can see that the rapid increase beyond the CF point is somewhat less steep for non T_1 states in these cases, but this is solely on the account of the CF point occurring at longer r due to presence of dynamic correlation.



(a) PBE (b) LRC-ωPBEh



(c) CIS

Figure 10.8: T_1 surfaces within the $M_S = \pm 1$ and $M_S = 0$ manifolds, as predicted by TDDFT/TDA. The $M_S = \pm 1$ states are spin pure (i.e. are purely green) while the $M_S = 0$ state shows spin-polarization past the CF point. The ground UKS/UHF S_0 state is also supplied for comparison.

The behavior for the spin-flipped block in TDDFT/TDA is interesting, on account of the lack of local exchange-correlation contributions to the f_{xc} kernel in the spin-flipped block. Fig 10.8 depicts the T_1 state obtained from the $M_S = \pm 1$ subspaces, along with the S_0 state and $M_S = 0$ T_1, for PBE, LRC-ωPBEh and HF/CIS. An important difference is evident prior to the onset of spin-polarization: the $M_S = \pm 1$ triplets predicted by the two density functionals are not degenerate with the $M_S = 0$ triplet, as should be the case in exact quantum mechanics (and as is the case in CIS till the CF point). The lack of degeneracy between the different subspaces in TDDFT/TDA even prior to spin-polarization in the reference is due to the lack of the local exchange-correlation contributions to f_{xc}.

Beyond the CF point, the TDDFT/TDA $M_S = \pm 1$ T_1 states start mimicking the behavior of their $M_S = 0$ counterpart, in sharp contrast to CIS. The $M_S = \pm 1$ T_1 surfaces display a sharp kink and rapid rise in energy to a dissociation limit that is non-degenerate with the S_0 state (as can be seen from Figs. 10.8a and 10.8b). The origin of this behavior can be understood from the form of Eqn 9.2. For a local functional like PBE, the f_{xc} contribution is zero. Simultaneously, i and a have different spins (as do j and b), voiding the $\langle ij|ab\rangle$ term. The spin-flip block of $\mathbf{A}$ therefore becomes a purely diagonal matrix of orbital energy differences, and all excitation energies correspond to those diagonal elements. The non-degeneracy between the α HOMO and the β LUMO (or vice versa) is nearly always guaranteed on account of the spin-polarized UKS potential they experience (even if the spatial orbitals are identical, as is the case for dissociation limit H_2). Consequently, the excitation energy cannot exactly become zero and the T_1 surface has to unphysically distort to accommodate this incorrect asymptotic behavior. The same general behavior applies to even hybrid functionals. Consequently, we conclude that the $M_S = \pm 1$ block yields physical T_1 surfaces only for CIS.

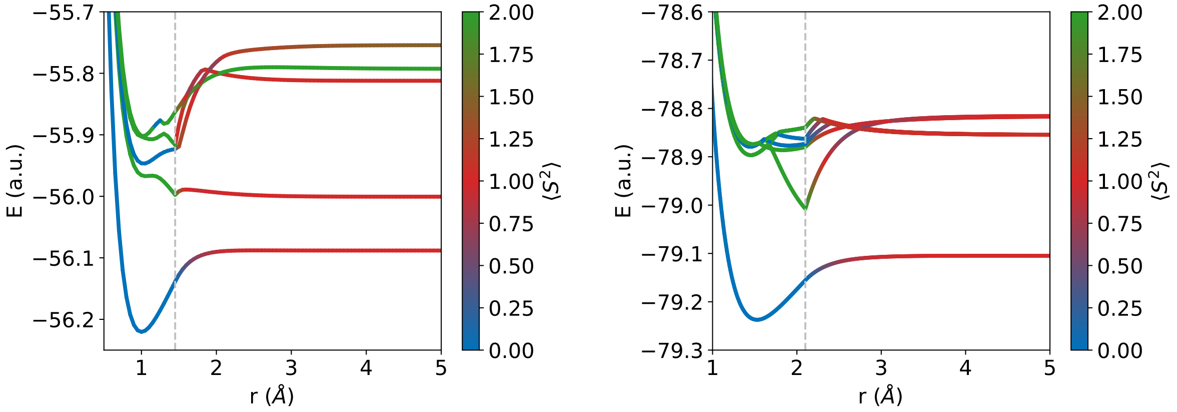
10.5 Examples for Other Systems

10.5.1 $M_S = 0$ subspace

Fig 10.9 shows the behavior of $M_S = 0$ UCIS for larger systems like the polar NH_3 and nonpolar ethane molecules. We see kinks appear at the CF point as well, along with dramatic jumps in many energy surfaces. The continuation of the T_1 state also does not go to the ground state dissociation limit. This indicates that our conclusions in the preceding subsection transfer to systems larger than H_2 as well, where more exact methods are no longer computationally affordable. TDDFT for larger species also exhibits similar behavior (with delocalization error driven lowering of CT state energies being the principal difference from CIS/TDHF) and some represenative PES are provided in the supporting information.

10.5.2 $M_S = \pm 1$ subspaces

CIS within the $M_S = \pm 1$ subspaces was effective in producing reasonable T_1 surfaces for H_2 (as can be seen from Fig 10.5b). This state of affairs however does not generalize to more complex systems, especially when polar bonds are involved. Let us now consider a



(a) NH_3/aug-cc-pVTZ along N-H stretch.

(b) C_2H_6/aug-cc-pVDZ along C-C stretch.

Figure 10.9: Ground and CIS excited states for larger species, within the M_S=0 subspace.

hypothetical system where an A–B bond is stretched, with the unpaired α electron localizing on fragment A while the unpaired β electron localizes on fragment B. To access the T_1 state with $M_S = 1$, we would need to flip the unpaired β electron on B to an α virtual orbital. The resulting state however would have a B fragment with $M_S = \frac{1}{2}$ but orbitals optimized for $M_S = -\frac{1}{2}$ configuration (i.e the unrestricted fragment ground state). Other spin-flip excitations can mimic some orbital rotation effects, but aside from the trivial case of B being H, the resulting state is thus higher in energy than the unrestricted ground state that has B with $M_S = -\frac{1}{2}$ and self-consistent orbitals, leading to the T_1 state going to an incorrect dissociation limit. Essentially, while there exists no energetic penalty for flipping a spin to go from $M_S = \pm\frac{1}{2}$ to $M_S = \mp\frac{1}{2}$ in exact quantum mechanics, the same is not true for CIS. Table 10.2 supplies some representative values of the error in the CIS spin-flip energies (which we call $\Delta_{\alpha\beta}$) for going from $M_S = \pm\frac{1}{2}$ to $M_S = \mp\frac{1}{2}$, which are small if the species in question is H like (i.e., alkali metals, that have one valence electron atop a noble gas core) but can be substantial if the unpaired electron has other occupied orbitals close to it in energy (like CH_3 or NH_2).

This spurious degeneracy error can manifest itself in three different manners for systems more complex than H_2 (assuming the direction of spin polarization is consistent).

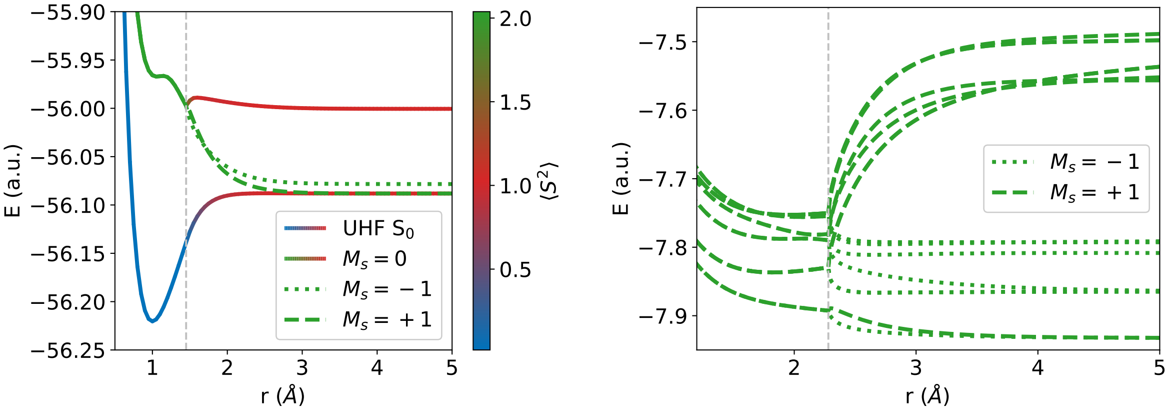
1. For A-H bond dissociations, the subspace including spin inversion on H will have a T_1 state that becomes degenerate with the unrestricted S_0 state. The other branch (which includes spin inversion on A) has a T_1 state that remains above the correct dissociation limit by $\Delta^A_{\alpha\beta}$. An example of this can be seen in Fig 10.10a, for NH_3.

2. For A-A bond dissociations, both subspaces will yield degenerate solutions on account of symmetry. However, both T_1 solutions will be above S_0 by $\Delta^A_{\alpha\beta}$ in the asymptotic limit.

Fragment	$\boldsymbol{\Delta_{\alpha\beta}}$
Li	0.0014
Na	0.0022
CH_3	0.2504
C_2H_5	0.2724
SiH_3	0.0884
NH_2	0.2702

Table 10.2: Degeneracy error (in eV) predicted by CIS for exciting from $M_S = \pm\frac{1}{2}$ to $M_S = \mp\frac{1}{2}$, for various radical fragments (A). This corresponds to the spurious degeneracy gap between the $M_S = \pm 1$ UCIS T_1 states at the dissociation limit of A-H bonds.

3. For A-B bond dissociations where B≠H, the two subspaces will yield different T_1 energies past the CF point, that remain distinct from the unrestricted ground state energy by $\Delta^A_{\alpha\beta}$ and $\Delta^B_{\alpha\beta}$ in the dissociation limit.

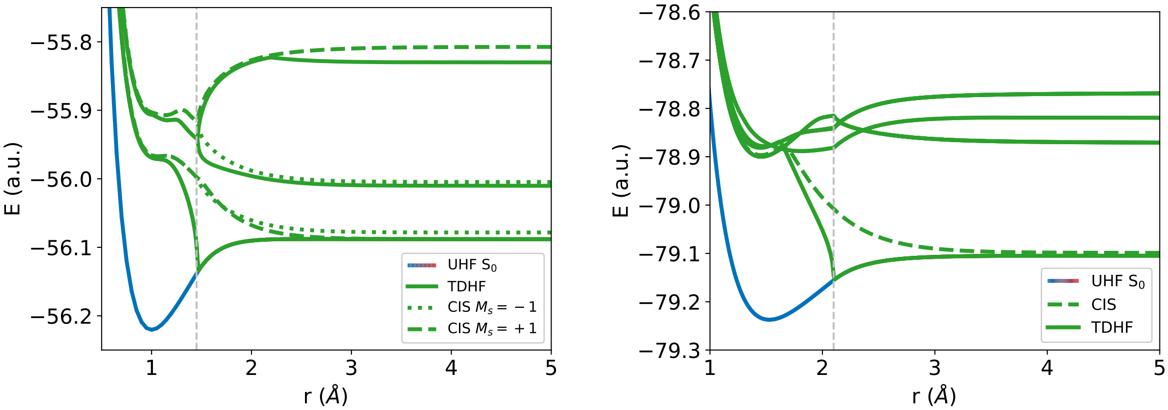


(a) NH_3/aug-cc-pVTZ. S_0 and $M_S = 0$ T_1 states are also depicted, for comparison.

(b) LiH/aug-cc-pVTZ.

Figure 10.10: T_1 predicted by CIS for spin-flip M_S. The spin-polarization consistently placed the down spin on the H, so only the $M_S = 1$ branch is asymptotically degenerate with the UHF ground state.

An additional feature for asymmetric bond dissociations is that the higher energy CIS solutions for the different M_S are only degenerate up to the CF point, and subsequently move apart, going to quite distinct roots of different character in the dissociation limit. Consider the LiH molecule (as depicted in Fig 10.10b). If we place the unpaired α electron on Li during spin-polarization, then excitations to the $M_S = -1$ subspace contain only local excitations on Li but not on H. Similarly, excitations to the $M_S = 1$ subspace capture local excitations on H , but not Li (barring very high energy core excitations). Similar considerations apply to CT transitions-it would be impossible to obtain a transfer of electron from H to Li in the $M_S = -1$ subspace, while only the core Li electron could be excited to H in the $M_S = 1$. The individual $M_S = \pm 1$ subspaces therefore contain complementary information for asymmetric bond dissociations.



(a) NH_3/aug-cc-pVTZ

(b) C_2H_6/aug-cc-pVDZ

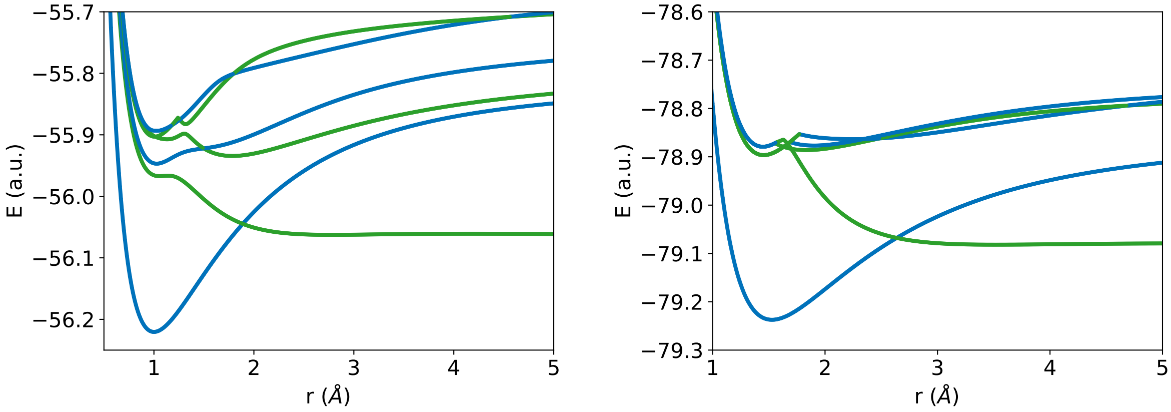
Figure 10.11: Triplet states predicted by TDHF for the spin-flip block. Some CIS states are provided for comparison, as is the UHF ground state.

The spin-flip TDHF solutions remain degenerate with the $M_S = 0$ solution till the CF point and then branch away. The T_1 states merge with the S_0 ground state beyond the CF point (as can be seen from Fig 10.11), yielding zero excitation energy—just like minimal basis H_2. The higher excited states however resemble the corresponding CIS states, and show similar branching behavior (as can be seen from Fig 10.11a).

An interesting side consequence of the spin-flip T_1 states merging with the S_0 state is that the $\Delta_{\alpha\beta}$ induced degeneracy error in CIS is completely absent in TDHF. Indeed, calculations on open-shell radical fragments (with $M_S = \frac{1}{2}$) show that $\Delta_{\alpha\beta} = 0$ in general. This can be viewed as a consequence of explicit orbital response terms contained within the **B** matrix. An alternative interpretation draws upon the GHF stability argument given earlier in Sec. IV, by noting that direction of the spin on the open-shell fragment is itself arbitrary within GHF, and so there should not be any energetic cost for rotation to a different spin direction.

10.5.3 RCIS

Figs 10.12a and 10.12b show RCIS excited state PESs for NH_3 and C_2H_6 respectively. The T_1 surfaces in both cases appear to reach reasonable dissociation limits, which are nonetheless considerably below the S_0 RHF dissociation limit (on account of spurious CT contributions



(a) NH_3/aug-cc-pVTZ along N-H stretch. (b) C_2H_6/aug-cc-pVDZ along C-C stretch.

Figure 10.12: RHF S_0 and RCIS excited states for larger species.

Bond	Asymptotic T_1 energy (in eV)
H—H	0.47 (0.47)
Li—H	0.60 (0.60)
Na—H	0.63 (0.63)
CH_3—CH_3	0.50 (0.67)
C_2H_5—H	0.78 (0.87)
SiH_3—H	0.54 (0.58)
NH_2—H	0.55 (0.62)

Table 10.3: Asymptotic T_1 energies (relative to ROHF independent fragments) predicted by RCIS for dissociation of A—B bonds. The T_1 energies relative to UHF fragments are given in parentheses.

in the latter). It is nonetheless worth noting that the asymptotic T_1 energies are above the dissociation limit of independent fragments by ≈ 0.5 eV (a more complete listing given in Table 10.3). The T_1 surfaces are also not monotonically decreasing (as is physically expected) but rather have small local minima with depths of ≈ 6 kJ/mol or so relative to the dissociation limits.

A comparison of Fig 10.12 with Fig 10.9 also shows that at longer separations, the RCIS surfaces are a lot less flat relative to the UCIS ones. This is likely on account of spurious CT character in the excited states, which is confirmed by examining the dipole moments of RCIS excited states of NH_3 at 50 Å. Aside from the T_1 state (which is nearly fully covalent), nine out of the ten lowest lying excited states have dipole moments in excess of 100 D, suggesting a fractional charge of ≈ 0.4 on fragments. Similar behavior is seen for LiH/aug-cc-pVTZ at 50 Å separation, indicating that this is not an unusual occurance for polar bonds. Interestingly, many excited state dipoles indicate charge transfer in the direction contrary to expectations based on electronegativity (i.e. $Li^{-\delta}H^{+\delta}$ vs the expected $Li^{+\delta}H^{-\delta}$), which is likely a consequence of CIS attempting to reverse the CT contamination in the RHF reference, but ultimately overcorrecting due lack of complete orbital response [81]. There is no net excited state dipole moment for low lying excited states of C_2H_6 because of the non-polar nature of the dissociating bond, but CT contributions are nonetheless present in both the reference and excited states, leading to incorrect asymptotic behavior of PESs. In general therefore, RCIS yields reasonable T_1 surfaces, but higher excited states often have substantial CT contamination on account of the RHF reference.

10.6 Conclusion

In conclusion, we have characterized TDDFT excited states (as well as those predicted by the related TDHF and CIS methods) for single bond dissociation, concentrating on the region beyond the Coulson-Fischer point where the ground state spin polarizes. We find that spin-polarization in the stable UKS (UHF) state beyond the CF point leads to disappearance of the initial $M_S = 0$ T_1 state at long internuclear separations. This is a consequence of the unrestricted excited state corresponding to the $M_S = 0$ T_1 having opposite spin-polarization than the ground state and is subsequently a double excitation away, despite being low lying in energy (and formally degenerate with the ground state at the dissociation limit). Kinks at the CF point are observed in many other excited state surfaces, along with spuriously elevated dissociation limits and unphysical curvature. It in fact appears that the spin-polarization zone just beyond the CF point connects what should naturally be two different surfaces via a rapidly increasing concave segment.

In CIS, triplet solutions within the spin-flip $M_S = \pm 1$ subspaces are non-degenerate with the standard $M_S = 0$ subspace beyond the CF point. The $M_S = \pm 1$ T_1 states roughly reach the correct dissociation limit for T_1. They are however non-degenerate past the CF point (or do not become exactly asymptotically degenerate with the ground S_0 state) due to insufficient orbital relaxation effects in CIS, for most single bond dissociation processes. The TDHF spin-flipped T_1 states are exactly degenerate with the UHF S_0 ground state beyond the CF point. These spin-flipping subspaces exhibit less desirable behavior when a functional with any local exchange-correlation contribution is employed. The shapes of the $M_S = \pm 1$ T_1 state surfaces predicted by typical TDDFT/TDA resemble the unphysical $M_S = 0$ T_1 surface.

We note that restricted CIS yields reasonable T_1 surfaces, despite the incorrect ground state dissociation limit solution. Of course, full TDDFT (or TDHF) on RHF/RKS solutions beyond the CF point is unwise on account of unphysical complex (or vanishing) excitation energies. Even for CIS, it must be noted that higher energy RCIS excited states, while smooth in contrast to their UCIS counterparts, tend to have significant amounts of spurious CT contamination, suggesting this approach has limited applicability and should be viewed with caution. If the spurious ionic terms in the ground state can be reduced, such as via approximate coupled cluster methods, the resulting restricted excited states will be more useful. Alternatively, some type of non-orthogonal CI [394, 395] could be employed to make the unrestricted methods useful, and the recently proposed holomorphic HF extensions [403] look promising in this regard.

Finally, it is important to keep in mind that processes more complex than single bond dissociation (like dissociation of O_2) could exhibit additional symmetry breaking (using complex or generalized orbitals) beyond the simple RKS→UKS instability discussed in this work. Such forms of symmetry breaking could induce even more exotic behavior in CIS/TDDFT excited state surfaces, which could prove even more challenging to address.

10.7 Computational Details

All calculations were performed with the Q-Chem 5.2 [191] package. Local exchange-correlation integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points for all atoms. All internal coordinates other than the stretch of the dissociating bond were held frozen at equilibrium configuration for polyatomic species (e.g. CH_3—CH_3 dissociates into unrelaxed trigonal pyramidal CH_3 radicals instead of relaxed, trigonal planar CH_3 radicals), for simplicity.

10.8 Appendix: Spin-flip Block TDHF Solutions for Minimal Basis H_2

It is possible to analytically show that the TDHF excitation energies for the spin-flipping $M_S = \pm 1$ blocks are zero for minimal basis H_2. We follow the treatment in Ref 14 and use the orbital labels employed in Table 10.1. Since there is only one occupied and one virtual orbital in both spin subspaces, the **A** and **B** matrices in the spin flip block are simply scalars. Mathematically, it implies:

$$\mathbf{A}_{\alpha\beta,\alpha\beta} = \left(\epsilon^{\bar{\beta}_2} - \epsilon^{\alpha_1}\right) + \left\langle \alpha_1\alpha_1 | \bar{\beta}_2\bar{\beta}_2 \right\rangle - \left\langle \alpha_1\bar{\beta}_2 | \alpha_1\bar{\beta}_2 \right\rangle \tag{10.9}$$

$$\mathbf{A}_{\beta\alpha,\beta\alpha} = \mathbf{A}_{\alpha\beta,\alpha\beta} \text{ (From spatial symmetry)} \tag{10.10}$$

$$\mathbf{B}_{\alpha\beta,\beta\alpha} = \left\langle \alpha_1\alpha_2 | \bar{\beta}_2\bar{\beta}_1 \right\rangle - \left\langle \alpha_1\bar{\beta}_2 | \alpha_2\bar{\beta}_1 \right\rangle = \mathbf{B}_{\beta\alpha,\alpha\beta} \tag{10.11}$$

However, the two electron integrals $\left\langle \alpha_1\alpha_1 | \bar{\beta}_2\bar{\beta}_2 \right\rangle$ and $\left\langle \alpha_1\alpha_2 | \bar{\beta}_2\bar{\beta}_1 \right\rangle$ are zero, as the spin parts integrate to zero.

Let us furthermore denote matrix elements of the Hamiltonian in the RHF basis (as in Ref 14), which gives us one electron matrix elements $h_{\sigma\sigma}$ and $h_{\sigma^*\sigma^*}$ and two electron matrix elements $\langle\sigma\sigma|\sigma\sigma\rangle = J_{\sigma\sigma}$ (self-repulsion in orbital $|\sigma\rangle$), $\langle\sigma^*\sigma^*|\sigma^*\sigma^*\rangle = J_{\sigma^*\sigma^*}$ (self-repulsion in orbital $|\sigma^*\rangle$), $\langle\sigma\sigma^*|\sigma\sigma^*\rangle = J_{\sigma\sigma^*}$ (repulsion between electrons in orbitals $|\sigma\rangle$ and $|\sigma^*\rangle$) and

$\langle \sigma\sigma^* | \sigma^*\sigma \rangle = K_{\sigma\sigma^*}$ (exchange interaction between electrons in orbitals $|\sigma\rangle$ and $|\sigma^*\rangle$). Other terms like $h_{\sigma\sigma^*}$ or $\langle \sigma\sigma | \sigma\sigma^* \rangle$ cancel out during the simplification and do not enter the picture.

With this notation, we have:

$$\mathbf{B}_{\alpha\beta,\beta\alpha} = (J_{\sigma\sigma} + J_{\sigma^*\sigma^*} - 2J_{\sigma\sigma^*})\cos^2\theta \sin^2\theta - \cos^2 2\theta K_{\sigma\sigma^*} \tag{10.12}$$

$$\mathbf{A}_{\alpha\beta,\alpha\beta} = \cos 2\theta \left(h_{\sigma^*\sigma^*} - h_{\sigma\sigma}\right) - \cos^4\theta J_{\sigma\sigma} - \sin^4\theta J_{\sigma^*\sigma^*} + \left(\cos^4\theta + \sin^4\theta\right) J_{\sigma\sigma^*} - \cos^2 2\theta K_{\sigma\sigma^*} \tag{10.13}$$

$$= \cos 2\theta \left(h_{\sigma^*\sigma^*} - h_{\sigma\sigma}\right) - \cos^2\theta \cos 2\theta J_{\sigma\sigma} - \cos^2\theta \sin^2\theta J_{\sigma\sigma} + \sin^2\theta \cos 2\theta J_{\sigma^*\sigma^*} - \sin^2\theta \cos^2\theta J_{\sigma^*\sigma^*} + \cos^2 2\theta J_{\sigma\sigma^*} + 2\sin^2\theta \cos^2\theta J_{\sigma\sigma^*} - \cos^2 2\theta K_{\sigma\sigma^*} \tag{10.14}$$

$$= \cos 2\theta \left(h_{\sigma^*\sigma^*} - h_{\sigma\sigma} - \cos^2\theta J_{\sigma\sigma} + \sin^2\theta J_{\sigma^*\sigma^*} + \cos 2\theta J_{\sigma\sigma^*} - 2\cos 2\theta K_{\sigma\sigma^*}\right) - \cos^2\theta \sin^2\theta J_{\sigma\sigma} - \sin^2\theta \cos^2\theta J_{\sigma^*\sigma^*} + 2\sin^2\theta \cos^2\theta J_{\sigma\sigma^*} + \cos^2 2\theta K_{\sigma\sigma^*} \tag{10.15}$$

However, the spin-polarized UHF solution satisfies $h_{\sigma^*\sigma^*} - h_{\sigma\sigma} - \cos^2\theta J_{\sigma\sigma} + \sin^2\theta J_{\sigma^*\sigma^*} + \cos 2\theta J_{\sigma\sigma^*} - 2\cos 2\theta K_{\sigma\sigma^*} = 0$ (Eqn. 3.374 in Ref 14), leading to:

$$\mathbf{A}_{\alpha\beta,\alpha\beta} = -\cos^2\theta \sin^2\theta J_{\sigma\sigma} - \sin^2\theta \cos^2\theta J_{\sigma^*\sigma^*} + 2\sin^2\theta \cos^2\theta J_{\sigma\sigma^*} + \cos^2 2\theta K_{\sigma\sigma^*} = -\mathbf{B}_{\alpha\beta,\beta\alpha} \tag{10.16}$$

The eigenvalues of $\mathbf{A}_{\alpha\beta,\alpha\beta}$ are the CIS excitations energies ω_{CIS} within the SF block. Consequently, Eqn. 10.6 simplifies to:

$$\begin{pmatrix} \omega_{CIS} & -\omega_{CIS} \\ -\omega_{CIS} & \omega_{CIS} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega_{TDHF} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{10.17}$$

$$\implies \begin{pmatrix} \omega_{CIS} & -\omega_{CIS} \\ \omega_{CIS} & -\omega_{CIS} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega_{TDHF} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{10.18}$$

whose only eigenvalue is 0, corresponding to the $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ eigenvector. Therefore, the TDHF excitation energies within the SF block are zero beyond the CF point, which is consistent with numerical observation.

10.9 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: TDDFT and TDDFT/TDA PES for LiH, NH_3 and C_2H_6 (pdf).

Chapter 11

Square Gradient Minimization

11.1 Introduction

Accurate quantum chemical methods for modeling electronic excited states are essential for gaining insight into the photophysics and photochemistry of molecules and materials. The most widely used technique for excited state calculations at present is time-dependent density functional theory (TDDFT) [250, 251, 306, 379, 389], on account of its relatively low computational complexity ($O(N^{2-3})$, where N is the molecule size) and reasonable accuracy for many problems [323, 324]. TDDFT excited states are computed via determining the linear-response of a ground state DFT solution to time-dependent external electric fields [251], permitting simultaneous modeling of multiple excited states. In principle, TDDFT is formally exact [250] when the exact time-dependent kernel of the exact exchange-correlation (xc) ground state functional is employed, although lack of that functional, and the need for the widely used adiabatic local density approximation [251, 379, 389] (ALDA) prevents this from being the case in practice. ALDA in fact restricts utility of TDDFT to single excitations out of the reference alone, with large errors arising whenever the target excited state has significant double (or higher) excitation character [325–327, 404]. Furthermore, TDDFT is known to systematically underestimate excitation energies for charge-transfer [251, 270, 387, 405] and Rydberg [325, 406] states (although long-range corrected functionals can be tuned to ameliorate these problems [330, 388]), and yields qualitatively erroneous potential energy surfaces along single bond dissociation coordinates [228]. These effects originally stem from errors in the ground state DFT solution like delocalization error [64, 95] or spin symmetry breaking [14, 218], but the linear-response protocol augments these deficiencies in the reference to catastrophic levels in excited states, on account of insufficient orbital relaxation [265, 321, 407, 408].

Orbital optimized excited state methods have consequently seen a renewal of interest in recent years [341, 351–354, 363, 369, 409–415], and have been successfully applied to problems like core excitations [376, 416] and CT states [265, 410] where orbital relaxation is expected to play a key role. However, excited state orbital optimization is fundamentally a challenging task due to the possibility of collapsing back into the ground state (often described as "variational collapse") as excited states are typically saddle point solutions of the orbital optimization equations. The Maximum Overlap Method (MOM) [341] attempts to mitigate this by selecting occupied orbitals after each iteration via maximization of overlap with the occupied orbitals from the previous iteration (instead of filling orbitals in ascending order of their energies). MOM has been quite successful in reducing the frequency of variational collapse, but has not fully eliminated it in practice [343, 344].

The continuing spectre of variational collapse has subsequently led to attempts to develop alternative variational principles [352, 353, 417] where excited states are true minima instead of saddle points. Such principles often employ the energy variance, which involves the $\mathbf{H}_e^2$ operator. The matrix elements of $\mathbf{H}_e^2$ are quite computationally challenging due to the presence of three and four particle operators. Furthermore, a straightforward generalization to DFT is not possible as xc functionals approximate $\langle \mathbf{H}_e \rangle$ and not $\langle \mathbf{H}_e^2 \rangle$, making it difficult to capitalize on the enormous strides made in Kohn-Sham DFT (KS-DFT) functional development for (ground state) energetics [55, 98, 111, 112] and properties [80, 226] in recent years.

In this work, we present a general approach that can extend *any* ground state orbital-optimization method to excited states, without any apparent onset of variational collapse. The computational cost of this approach is only about 3 times the cost of ground state orbital optimization (per iteration), when implemented via a simple finite-difference protocol based on analytic orbital gradients of the energy/Lagrangian. This method is subsequently applied to two excited state orbital optimized DFT techniques: ΔSCF and Restricted Open-Shell Kohn-Sham (ROKS). The utility of these approaches is demonstrated via application to theoretically well characterized double excitations of small molecules, singly excited states of formaldehyde and an analysis of the absorption spectrum of zinc phthalocyanine.

11.2 Theory

11.2.1 Objective Function

Orbital optimization based methods attempt to minimize some Lagrangian $\mathcal{L}$ against orbital degrees of freedom $\vec{\theta}$ (that mix occupied orbitals $\{i\}$ with virtual orbitals $\{a\}$). $\mathcal{L}$ is simply the energy for Hartree-Fock (HF) or DFT (or indeed, any variational method), although it can be considerably more complex (eg. including amplitude constraint terms) for non-variational methods like Møller-Plesset perturbation theory [266, 267, 276, 418] or coupled cluster. [419] For excited states, the typical objective is to find an unstable extremum of $\mathcal{L}$ instead of the global minimum, which is quite challenging due to the possibility of variational collapse down to a minimum. While (quasi) Newton methods can be successful in converging to a saddle point if supplied with an excellent initial orbital guess and good initial $\mathcal{L}$ Hessian, the possibility of collapsing into a local minimum instead remains fairly high.

We convert the extremization problem into a minimization by instead focusing on:

$$\Delta = \left|\nabla_{\vec{\theta}}\mathcal{L}\right|^2 = \sum_{ai}\left(\frac{\partial \mathcal{L}}{\partial \theta_{ai}}\right)^2 \tag{11.1}$$

Δ therefore is merely the square of the gradient of $\mathcal{L}$ against orbital degrees of freedom $\vec{\theta}$, and is therefore positive semidefinite by construction. $\Delta = 0$ if and only if $\nabla_{\vec{\theta}}\mathcal{L} = 0$, which indicates stationarity of $\mathcal{L}$ against the orbital degrees of freedom. The challenges typically encountered in optimizing unstable extrema (i.e. saddle points or maxima) in $\mathcal{L}$ are therefore averted, as every orbital optimized state is a global minima of Δ. Other extrema are possible, as discussed later, but are easily identifiable by $\Delta \neq 0$.

11.2.2 Gradient

The gradient of Δ with respect to $\vec{\theta}$ is given by:

$$\frac{\partial}{\partial \theta_{ai}}\Delta = \frac{\partial}{\partial \theta_{ai}}\sum_{bj}\left(\frac{\partial \mathcal{L}}{\partial \theta_{bj}}\right)^2 = 2\sum_{bj}\left(\frac{\partial \mathcal{L}}{\partial \theta_{bj}}\right)\left(\frac{\partial^2 \mathcal{L}}{\partial \theta_{bj}\partial \theta_{ai}}\right) \tag{11.2}$$

For HF/DFT, the cost of analytically evaluating the gradient via the matrix-vector contraction $\sum\limits_{bj}\left(\frac{\partial \mathcal{L}}{\partial \theta_{bj}}\right)\left(\frac{\partial^2 \mathcal{L}}{\partial \theta_{bj}\partial \theta_{ai}}\right)$ should roughly equal the cost of constructing the Fock matrix $\mathbf{F}$. The cost of analytically evaluating $\nabla_{\vec{\theta}}\Delta$ should therefore be twice the cost of evaluating $\nabla_{\vec{\theta}}\mathcal{L}$: once for constructing $\nabla_{\vec{\theta}}\mathcal{L}$, and another for the contraction with the Hessian $\left(\frac{\partial^2 \mathcal{L}}{\partial \vec{\theta}\partial \vec{\theta}'}\right)$. Analytical Δ gradients are therefore straightforward at the HF/DFT level, at a compute cost of roughly twice the analytical $\mathcal{L}$ orbital gradient. However, efficient implementation of the analytic $\mathcal{L}$ Hessian is undoubtedly challenging for more complex methods.

A simple finite difference approach permits us to sidestep this issue for very little additional cost. Such an approach has already been used for orbital stability analysis (i.e. extremal eigenvalues of the orbital Hessian) [276, 420], building on earlier work evaluating extremal eigenvalues of the force constant matrix (Hessian with respect to nuclear displacements). [421–424] We know that:

$$\left(\frac{\partial \mathcal{L}}{\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}+\vec{\delta\theta}} = \left(\frac{\partial \mathcal{L}}{\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}} + \sum_{bj}\left(\frac{\partial^2 \mathcal{L}}{\partial \theta_{bj}\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}}\delta\theta_{bj} + O((\vec{\delta\theta})^2) \tag{11.3}$$

from a Taylor expansion of the derivative $\frac{\partial \mathcal{L}}{\partial \theta_{ai}}$ about the point $\vec{\theta} = \vec{\theta_0}$, on account of a perturbation $\vec{\delta\theta}$. Subsequently, we can choose $\vec{\delta\theta} = \lambda\left(\nabla_{\vec{\theta}}\mathcal{L}\right)_{\vec{\theta}=\vec{\theta_0}}$, which yields:

$$\sum_{bj}\left(\frac{\partial \mathcal{L}}{\partial \theta_{bj}}\right)_{\vec{\theta}=\vec{\theta_0}}\left(\frac{\partial^2 \mathcal{L}}{\partial \theta_{bj}\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}} = \frac{1}{2\lambda}\left(\left(\frac{\partial \mathcal{L}}{\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}+\vec{\delta\theta}} - \left(\frac{\partial \mathcal{L}}{\partial \theta_{ai}}\right)_{\vec{\theta}=\vec{\theta_0}-\vec{\delta\theta}}\right) + O\left(\lambda^2\right) \tag{11.4}$$

$$\implies \left(\nabla_{\vec{\theta}}\Delta\right)_{\vec{\theta}=\vec{\theta_0}} = \frac{1}{\lambda}\left(\left(\nabla_{\vec{\theta}}\mathcal{L}\right)_{\vec{\theta}=\vec{\theta_0}+\vec{\delta\theta}} - \left(\nabla_{\vec{\theta}}\mathcal{L}\right)_{\vec{\theta}=\vec{\theta_0}-\vec{\delta\theta}}\right) + O\left(\lambda^2\right) \tag{11.5}$$

In other words, taking the finite difference between the gradient $\nabla_{\vec{\theta}}\mathcal{L}$ at two slightly shifted $\vec{\theta}$ (with the shift being proportional to the gradient $\nabla_{\vec{\theta}}\mathcal{L}$ at the central point) yields the desired Hessian-gradient contraction. The cost of this approach for finding $\nabla_{\vec{\theta}}\Delta$ is therefore thrice the cost of a single $\nabla_{\vec{\theta}}\mathcal{L}$ gradient evaluation, which is not a substantial increase over the 2× cost associated with contraction with the analytic Hessian. While this approach does introduce precision errors associated with finite differencing, their magnitude can be controlled via judicious choice of λ. More importantly, the errors scale as $O\left(\lambda^2\left(\left(\nabla_{\vec{\theta}}\mathcal{L}\right)^*\right)^3\right)$, indicating that they are the largest when we are far from convergence (i.e. large $\nabla_{\vec{\theta}}\mathcal{L}$) when a very accurate gradient is not critical. The errors should be quite small close to convergence (when $\nabla_{\vec{\theta}}\mathcal{L}$ should be small). Alternative higher order finite difference formulae could also be employed, though we shall not consider such choices here.

11.2.3 Preconditioner

The convergence of a gradient based optimization process can be dramatically accelerated via use of appropriate preconditioners, like a diagonal approximation to the Hessian. This is especially true for a relatively less well conditioned problem like Δ minimization. Unfortunately, exact evaluation of the diagonal terms of the Δ Hessian is likely far too computationally demanding to be worthwhile. However, mean-field terms (i.e. $\mathbf{F}$ terms) make up the largest portion of $\mathcal{L}$ for non-strongly correlated species. Focusing on those terms alone suggests that within a pseudocanonical orbital basis (i.e. occupied-occupied and virtual-virtual blocks of $\mathbf{F}$ are diagonal), an approximate preconditioner $B_{ia,jb} = 8\left(\epsilon_a - \epsilon_i\right)^2\delta_{ia}\delta_{jb}$ (where ϵ_a and ϵ_i are energies of pseudocanonical spin-orbitals a and i, respectively) would be appropriate. This is basically a generalization of the preconditioner used in the geometric direct minimization (GDM) method [348] for ground state minimization.

11.2.4 The SGM Method for Orbital Optimization

The gradient and the preconditioner described in the two preceding subsections can be employed to minimize Δ, starting from any initial guess in orbital space. To do so, we build upon the GDM quasi-Newton method [425] for $\mathcal{L}$ minimization (which uses the BFGS update [426–429]). For our squared gradient minimization (SGM) problem, we supply the gradient and preconditioner appropriate for Δ to the GDM algorithm. The computational cost of a single SGM iteration should therefore be at most three times the cost of the corresponding GDM iteration for the ground state.

SGM would ideally converge to the closest solution in orbital space, when supplied with an initial set of guess orbitals. In practice however, the highly approximate nature of the initial preconditioner could result in large initial steps that lead to convergence to an alternative root (or even the ground state!). However, this can be easily mitigated by scaling the gradient down by some scalar c to minimize the size of the initial steps in order to prevent large initial stepsizes. The approximate BFGS Hessian would however be effectively scaled by the same c, and the long term convergence rate not be (too) negatively impacted. We have found that $c = 1$ is typically adequate for most cases, but a very low value of $c = 0.01$ could be employed as a conservative choice for difficult cases. We also note that Δ minimization is less well conditioned than energy/Lagrangian minimization (due to squaring of the gradient) and SGM is consequently an inefficient ground state optimizer (which can also converge to unwanted saddles if starting from a poor initial guess), relative to conventional GDM.

11.2.5 Relationship with Other Methods

Δ minimization via SGM is essentially a generalization of GDM [348] for saddle point optimization. It is consequently a direct minimization based alternative to $\mathbf{F}$ matrix diagonalization methods like the Maximum Overlap Methods (MOM) [341, 344]. In addition, SGM is closely related to the excited state variational principle employed by the Excited State Mean-Field (ESMF) approach described in Ref 352 and the σ-SCF approach described in Ref 353. The objective function in Ref 352 reduces to Δ if its energy targeting component is deleted. This has very recently been generalized to a generalized variational principle (GVP) [369], which smoothly scales various components, and can thus become Δ in some limits. The presence of the energy targeting term enables GVP to target states close to a particular input energy, while SGM Δ minimization aims to converge to the closest minima in orbital space to the initial guess. SGM is therefore simpler, though necessarily more guess-dependent. The more general form of the GVP also permits it to switch from minimization to saddle point search, which however comes with some risk of collapse if the minimization component has not adequately converged a good starting point for the saddle point convergence. SGM has no such issues as it is a pure minimization. The use of Tensorflow's [430] automatic differentiation package in Ref 369 also restricts applicability to large systems, while our analytic/finite-difference based Δ gradients do not have such issues. It is also worth noting that a finite difference approach was employed in Ref 369 for Newton-Raphson iterations, although the per-iteration computational cost was larger due to the need to construct a large Krylov subspace for inverting the Hessian of the objective function.

The parallels with σ-SCF [353] are less obvious at first glance. σ-SCF minimizes the energy variance $\sigma^2 = \left\langle \Phi \left| (\mathbf{H}_e - \langle \mathbf{H}_e \rangle)^2 \right| \Phi \right\rangle = \langle \Phi | \mathbf{H}_e^2 | \Phi \rangle - (\langle \Phi | \mathbf{H}_e | \Phi \rangle)^2$ for a single determinant $|\Phi\rangle$. The computational expense of evaluation of $\langle \Phi | \mathbf{H}_e^2 | \Phi \rangle$ makes it natural to wonder if substitution of $\mathbf{H}_e$ with a mean-field 1 body Hamiltonian like the Fock-matrix $\mathbf{F}$ would be acceptable. The most challenging term would then be:

$$\left\langle \Phi \left| \hat{\mathbf{F}^2} \right| \Phi \right\rangle = \sum_{|D\rangle} \langle \Phi | \mathbf{F} | D \rangle \langle D | \mathbf{F} | \Phi \rangle \tag{11.6}$$

by doing a resolution of the identity over all determinants $|D\rangle$ in Hilbert space. $\langle \Phi | \mathbf{F} | D \rangle \neq 0$ only if $|D\rangle$ is either $|\Phi\rangle$ or a single excitation $|\Phi_i^a\rangle$. Therefore:

$$\left\langle \Phi \left| \hat{\mathbf{F}^2} \right| \Phi \right\rangle = \left(\langle \Phi | \mathbf{F} | \Phi \rangle\right)^2 + \sum_{i,a} \langle \Phi | \mathbf{F} | \Phi_i^a \rangle \langle \Phi_i^a | \mathbf{F} | \Phi \rangle \tag{11.7}$$

$$\implies \left\langle \Phi \left| \hat{\mathbf{F}^2} \right| \Phi \right\rangle - \left(\langle \Phi | \mathbf{F} | \Phi \rangle\right)^2 = \sum_{i,a} \langle \Phi | \mathbf{F} | \Phi_i^a \rangle \langle \Phi_i^a | \mathbf{F} | \Phi \rangle = \sum_{i,a} |F_{ai}|^2 = \frac{1}{4}\Delta \tag{11.8}$$

In essence, SGM (or any single determinant optimizer like MOM) performs mean-field variance minimization, in contrast to the full $\hat{\mathbf{H}}$ based variance minimization of σ-SCF.

11.2.6 Local Extrema in Δ

From Eqn 11.2, we can infer that $\nabla_{\vec{\theta}}\Delta = 0$ implies either $\nabla_{\vec{\theta}}\mathcal{L} = 0$ (indicating successful extremization) or that the gradient $\nabla_{\vec{\theta}}\mathcal{L}$ belongs to the null-space of the Hessian $\dfrac{\partial^2 \mathcal{L}}{\partial\vec{\theta}\partial\vec{\theta}'}$. While cases with singular $\dfrac{\partial^2 \mathcal{L}}{\partial\vec{\theta}\partial\vec{\theta}'}$ are known in quantum chemistry as Coulson-Fischer points [212], such points are defined by zero gradients. Hence little or nothing is known about singular orbital hessians with *nonzero* gradients. Interestingly, we have encountered a few such solutions over the course of our investigations, as indicated by $\Delta \neq 0$ at convergence. We were able to escape them via use of "better" initial guesses, such as by providing converged local spin-density approximation (LSDA) excited state orbitals to a hybrid DFT calculation (instead of ground state hybrid DFT orbitals).

11.3 Excited State Orbital Optimized DFT

11.3.1 ΔSCF

ΔSCF [264, 358] methods converge a single Slater determinant as an excited state solution to the HF/KS equations. The likelihood of variational collapse had long restricted the utility of ΔSCF, but the development of MOM led to a revival of interest in the method [341, 358, 416, 431]. MOM nonetheless does not always succeed in averting variational collapse (as will be shown later), making it desirable to have alternative solvers for challenging cases.

Apart from convergence, other main concerns with ΔSCF are twofold. The Hohenberg-Kohn theorem [41] does not formally hold for excited states [367], and it cannot be assumed that ground state functionals will be accurate for excited states. ΔSCF is thus a pragmatic choice for modeling excited states in large systems, but will not be a foolproof solution. Nonetheless, practical studies have shown that quite high levels of accuracy can be obtained from ΔSCF [341, 344, 358, 416, 431] for challenging problems that TDDFT fails to address properly, without compromising accuracy in TDDFT's ideal domain of applicability (valence excitations in closed shell species). Our results also demonstrate this, as will be shown later.

The second, closely related, challenge facing ΔSCF is that unlike ground states, excited states cannot often be well approximated by a single Slater determinant. In particular, single excitations out of closed-shell molecules are intrinsically multiconfigurational, as both α and β electrons are equally likely to be excited (leading to at least two configurations of equal weights). ΔSCF within the $M_S = 0$ subspace can only target one of the configurations and would therefore yield a heavily spin-contaminated ("mixed") determinant with $\langle S^2 \rangle \approx 1$ for even otherwise well-behaved single excitations. This is an issue for singlet excited states alone, as $M_S = \pm 1$ triplet states are typically well described by single determinants and thus ΔSCF. The singlet energies can be approximated via approximate spin-purification [357], if the only major spin-contaminant is the corresponding triplet. It would however be ideal to orbital optimize the spin-purified energy directly instead of optimizing the mixed and triplet configurations separately. The Restricted Open-Shell Kohn-Sham (ROKS) method achieves this for pure open shell singlet states, and is described in the next subsection. Other potentially more general alternatives like half-projection [354] or the DFT generalization to ESMF [415] appear to possess double counting errors, making them somewhat less appealing.

11.3.2 ROKS

The Restricted Open-Shell Kohn Sham (ROKS) technique [362, 363] optimizes orbitals for spin-pure singly excited states by extremizing

$$\mathcal{L}_{\text{ROKS}} = 2E_M - E_T \tag{11.9}$$

where E_M is the energy of the mixed determinant and E_T is the energy of the triplet within the $M_S = 1$ subspace, using the same spin-restricted orbitals. This is reasonable for true open shell singlets, as the mixed state is half singlet and half triplet when RO orbitals are used. The same strategy could also be applied to double excitations where a single electron pair has been broken, such as the ${}^1B_{3g}$ state of tetrazine [355]. ROKS has been shown to be quite effective at predicting HOMO→LUMO type excitations in organic molecules [363] and is excellent for CT state energies in systems where TDDFT fails catastrophically [265]. However, the implementation described in Ref 363 is restricted to the lowest excited singlet (S_1) state alone. SGM however permits application of ROKS to arbitrary excited singlet states without collapse back to the S_1 state, thus considerably generalizing its applicability to excited state calculations. Of course, ROKS is itself limited in applicability: ROKS can only describe transitions that are well represented as promotions from one spatial occupied orbital to one spatial virtual orbital after orbital optimization. Excitations that can only be represented by transitions between multiple orbital pairs that have no common orbitals are unlikely to be well described on account of their natively multiconfigurational nature. A rather well-known example of the latter are the L_b dark states in polyaromatic compounds [432].

11.3.3 Constrained Methods

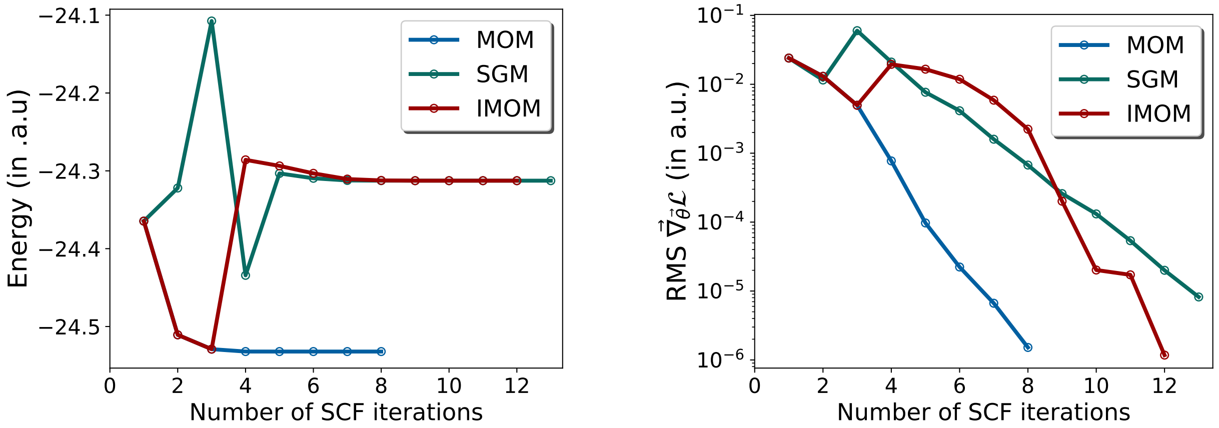
There exist other excited state DFT ansatze that employ constraints for the purpose of transferring electrons into the virtual space of the ground state determinant or attempt to impose strict orthogonality between ground and excited state determinants, for the purpose of avoiding collapse back to the ground state. Some examples of such methods are SCF constrained variational DFT (SCF-CV-DFT) [433, 434], orthogonality constrained DFT (OCDFT) [435] or excited constrained DFT (XCDFT) [436]. SGM could in principle be applied to these as well, although we have not attempted to do so at present.

11.4 Applications

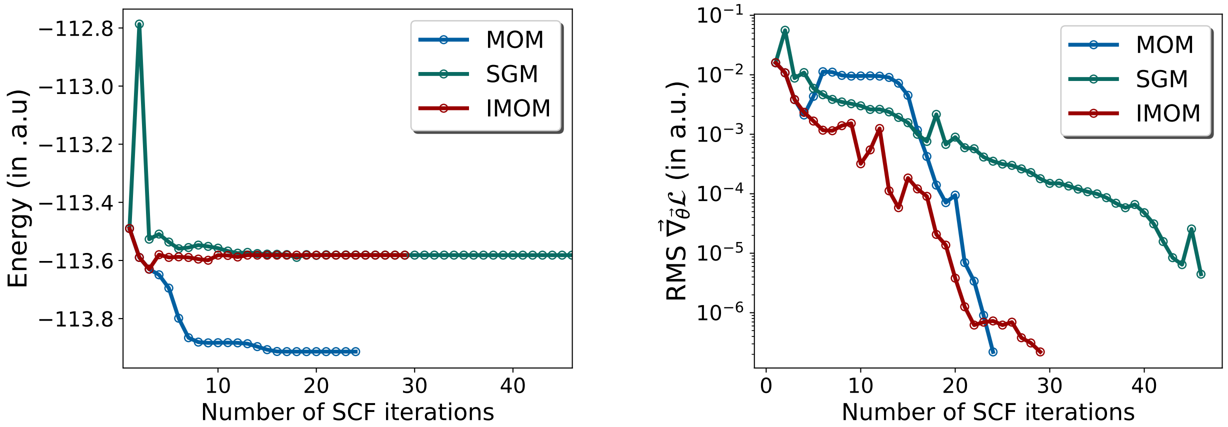
11.4.1 Comparison of SGM to MOM and IMOM

MOM has encountered considerable success in averting variational collapse for ΔSCF, but is nonetheless not a perfect solution [343]. Two systems where MOM fails to avert variational collapse are the 2p→ 3p excitation in the B atom [344] and a Rydberg-like single excitation out of the highest energy oxygen lone-pair to a C 4p_y orbital in formaldehyde. SGM however is successful at converging both, as can be seen from the plots in Fig 11.1. As shown in Fig 11.1, the Initial MOM (IMOM) method [344] (which selects occupied orbitals at the end of a **F** diagonalization based on overlap with an *initial* set of orbitals vs the ones from the *preceding* step) is also able to converge to the same solution as SGM for both of these cases, for a considerably smaller computational cost (stemming from fewer **F** builds being required).

IMOM's good performance stems from it avoiding 'drifting' of orbitals away from the initial guess over multiple SCF iterations. However, IMOM can exhibit oscillatory behavior and fail to converge to a solution, as exhibited by the case of an excitation from the highest energy π lone pair to the second lowest π^* orbital in nitrobenzene (as depicted in Fig 11.2), where IMOM shows no sign of convergence even after 500 DIIS [342] steps. This is likely a consequence of IMOM picking significantly different orbitals after some steps, based on discontinous ranking changes arising from small fluctuations in the overlap with the initial orbitals. On the other hand, MOM monotonically collapses back to the ground state. Considering all three examples, SGM appears to be a relatively stable ΔSCF solver that could prove effective in converging challenging states, although it is more expensive than MOM or IMOM per iteration (and typically requires more steps when those methods converge).



(a) 2p→3p of B atom (UHF/aug-cc-pVTZ [198, 437]).



(b) O lone pair → C $4p_y$ Rydberg state of HCHO (UHF/aug-cc-pVTZ [198, 437]).

Figure 11.1: Energy and gradient ($\vec{\nabla}_{\vec{\theta}}\mathcal{L}$) convergence to ΔSCF solutions with MOM, IMOM and SGM. SGM converges energies to 10^{-8} a.u. in 13 iterations (39 Fock builds) for the B atom and 46 iterations (138 Fock builds) for HCHO, with $c = 1$. In contrast, IMOM requires 12 and 29 SCF cycles, respectively.

11.4.2 Application to Doubly Excited States

Doubly excited states (or states with significant double excitation character) are typically inaccessible via TDDFT due to use of the ALDA. The efficacy of ΔSCF for modeling such double excitations has already been hinted at [344, 431], leading us to study the extent to which ground state DFT functionals could reproduce vertical excitation energies for a few systems with theoretically well-characterized pure double excitations from Ref 355. ΔSCF solutions were examined for all states aside from the ${}^1B_{3g}$ state of tetrazine, where the presence of a broken electron pair necessitated use of ROKS (the corresponding doubly excited triplet was modelled with RO orbitals as well, to ensure spin-purity). SGM was used

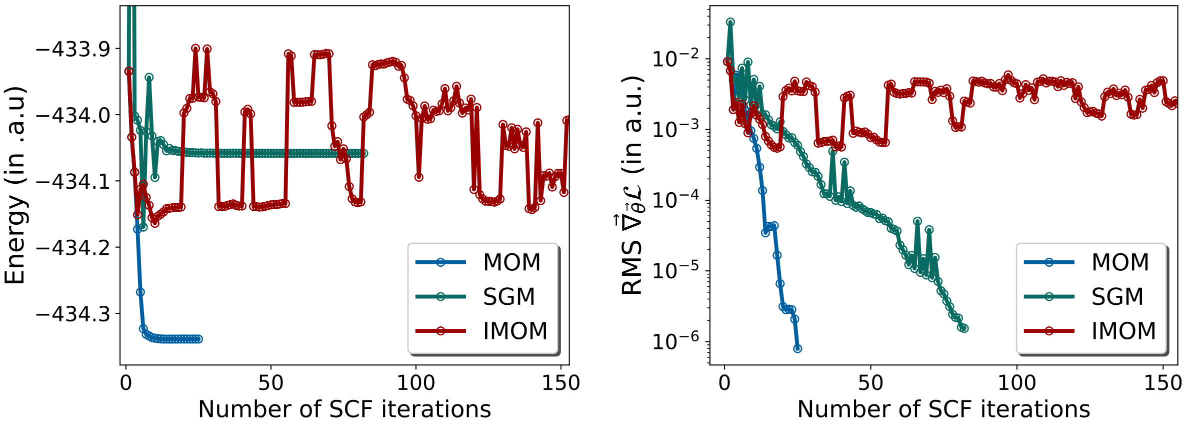


Figure 11.2: Energy and gradient ($\vec{\nabla}_{\vec{\theta}}\mathcal{L}$) convergence to ΔSCF solutions for the excitation from the highest energy π lone pair to the second lowest π^* orbital in nitrobenzene (UHF/def2-TZVP [438]). MOM collapses back to the ground state while IMOM fails to converge even after 500 iterations (only the first 150 are shown). SGM (with $c = 1$) however converges the energy to 10^{-8} a.u. after 82 iterations (246 Fock builds).

to converge all solutions.

The results shown in Table 11.1 demonstrate that orbital optimization with standard density functionals can achieve very good accuracy for the doubly excited states considered here, surpassing considerably more expensive wavefunction techniques like CC3 [439] that do not incorporate excited state orbital relaxation. Indeed, even the humble SPW92 [138, 149] LSDA functional (that is only accurate for the uniform electron gas) has a lower root mean squared error (RMSE) than CC3! It appears that the meta generalized gradient approximations (mGGAs) SCAN [62] and B97M-V [63] from Rung 3 of Jacob's ladder are very accurate for the double excitations studied, yielding rather small RMSEs ≤ 0.25 eV. The PBE0 [91] global hybrid GGA also performs well, with an RMSE of only 0.31 eV, while the range-separated hybrid, ωB97X-V [108] yields rather disappointing performance in light of its good accuracy for ground state energetics [55] and properties [80, 226].

The origin of this behavior could be partly understood by looking at the sensitivity of the predictions to the functional choice. The majority of the species in Table 11.1 show remarkably little functional sensitivity (with CH_3NO having a standard deviation of only 0.07 eV between predictions), but the lone pair to π^* transitions of glyoxal, pyrazine and tetrazine show significant sensitivity to the choice of functional (standard deviation of ≥ 0.5 eV). We therefore classify these species into a "sensitive" subset and the remainder into a "insensitive" one, with the subset RMSEs also reported in Table 11.1. SCAN, B97M-V, PBE0 and ωB97X-V give excellent (and very similar) performance for the five insensitive transitions, while LSDA and PBE perform somewhat more poorly. On the other hand, the

Species (Excitation)	Rung 1	Rung 2	Rung 3		Rung 4		CC3	TBE
	SPW92	PBE	SCAN	B97M-V	PBE0	ωB97X-V		
Be (2s$^2 \to$ 2p^2)	6.97	6.98	7.11	7.08	7.23	7.52	7.08	7.06
HNO (n$^2 \to (\pi^*)^2$)	4.00	4.13	4.24	4.33	4.24	4.26	5.21	4.32
HCHO (n$^2 \to (\pi^*)^2$)	9.56	9.73	10.02	10.06	10.07	10.20	11.18	10.34
C_2H_4 ($\pi^2 \to (\pi^*)^2$)	11.78	11.75	12.22	12.23	12.27	12.57	12.80	12.56
CH_3NO (n$^2 \to (\pi^*)^2$)	4.63	4.63	4.71	4.81	4.70	4.69	5.73	4.74
Glyoxal (n$^2 \to (\pi^*)^2$)	4.83	4.97	5.37	5.56	5.88	6.56	6.76	5.54
Pyrazine (n$^2 \to (\pi^*)^2$)	7.35	7.49	7.90	8.15	8.43	8.78	9.17	8.04
Tetrazine (n$^2 \to (\pi^*)^2$, ^1A_g)	3.99	4.14	4.55	4.89	5.10	5.72	6.18	4.60
Tetrazine(n$^2 \to \pi_1^*\pi_2^*$, $^3B_{3g}$)	4.77	4.87	5.24	5.73	5.92	6.78	7.34	5.51
Tetrazine(n$^2 \to \pi_1^*\pi_2^*$, $^1B_{3g}$)	5.12	5.24	5.62	6.19	6.40	7.20	7.60	6.14
RMSE	0.65	0.56	0.25	0.18	0.31	0.76	1.15	
RMSE (insensitive)	0.52	0.47	0.21	0.20	0.20	0.22	0.71	
RMSE (sensitive)	0.77	0.64	0.28	0.17	0.39	1.05	1.46	
ME	-0.58	-0.49	-0.19	0.02	0.14	0.54	1.02	
MAX	1.02	0.90	0.52	0.33	0.50	1.27	1.82	

Table 11.1: Vertical excitation energies (in eV) for pure double excitations with DFT/aug-cc-pVTZ [198, 200, 437], compared to CC3 and theoretical best estimates (TBE) from Ref 355. CC3 values from Ref 355 have been extrapolated to the complete basis set (CBS) limit. Root mean squared error (RMSE), mean error (ME) and maximum absolute error (MAX) relative to TBE are also reported.

sensitive transitions are considerably more challenging, with LSDA and PBE significantly underestimating the excitation energy, while ωB97X-V significantly overestimates it. SCAN, B97M-V and PBE0 make predictions intermediate to the two extremes and consequently have low error. These trends seem to correlate well with the delocalization error present in these functionals [95], although the significant difference in performance between SPW92/PBE and SCAN/B97M-V cannot be fully explained by any delocalization based argument alone. We do however note that a similar performance gap between Rung 1-2 and Rung 3 functionals were seen for static polarizability predictions [226], which is a global metric for accuracy of symmetry allowed singly excited states.

Despite these limitations, all functionals tested are more accurate than the $O(N^7)$ scaling CC3 method, at only $O(N^{3-4})$ cost. The very good performance of Rung 3 functionals is encouraging in this light, as it shows that useful results can be obtained from relatively inexpensive local functionals, permitting reasonable estimate of energies of doubly excited states for very large molecular systems or even extended materials.

	Transition (Symmetry)	SPW2	PBE	B97M-V	SCAN	PBE0	ωB97X-V	TBE
Singlet								
Valence	$n \to \pi^*$ ($^1A_2$)	3.81	3.65	3.84	3.51	3.62	3.81	3.97
	$\sigma \to \pi^*$ ($^1B_1$)	8.84	8.58	8.86	8.40	8.64	8.83	9.21
	$\pi \to \pi^*$ ($^1A_1$)	8.72	8.88	9.59	9.55	9.78	9.86	9.26
Rydberg	$n \to 3s$ ($^1B_2$)	7.02	6.92	7.11	6.99	7.06	7.30	7.3
	$n \to 3p$ ($^1B_2$)	7.83	7.71	8.00	7.83	7.89	8.23	8.14
	$n \to 3p$ ($^1A_1$)	7.87	7.71	7.98	7.80	7.89	8.25	8.27
	$n \to 3p$ ($^1A_2$)	8.36	8.13	8.58	8.26	8.31	8.73	8.50
Triplet								
Valence	$n \to \pi^*$ ($^3A_2$)	3.32	3.29	3.37	3.12	3.26	3.45	3.58
	$\pi \to \pi^*$ ($^3A_1$)	6.58	6.22	6.02	5.80	5.84	6.08	6.07
Rydberg	$n \to 3s$ ($^3B_2$)	6.84	6.73	6.83	6.81	6.91	7.21	7.14
	$n \to 3p$ ($^3B_2$)	7.66	7.55	7.70	7.68	7.74	8.12	7.96
	$n \to 3p$ ($^3A_1$)	7.74	7.58	7.78	7.72	7.79	8.19	8.15

Table 11.2: ROKS singlet excitation energies and RO-ΔSCF triplet excitation energies for HCHO in eV (using the aug-cc-pVTZ basis). The best theoretical estimates (TBE) has been obtained from Ref 440.

11.4.3 Singly Excited States of Formaldehyde

HCHO is a small molecule whose lowest lying excited states have been very well theoretically characterized [440], making it an ideal candidate for applying the SGM approach to converge ROKS for higher singlet excited states. The resulting excitation energies are shown in Table 11.2, along with ΔSCF energies for the triplet state within the $M_S = 1$ subspace (using RO orbitals for consistency with ROKS). Corresponding TDDFT numbers have been provided in the Supporting Information. Quantitative errors for all methods (along with corresponding values from TDDFT and some wavefunction theories) are reported in Table 11.3.

The values in Table 11.3 stem from only one species, but nonetheless contain some inferences that are likely to be transferable. First, using ROKS for singlet excited states (and RO-ΔSCF for triplets) does not lead to any degradation in performance for valence excitations, consistent with previous studies [358, 363]. This is unsurprising, as standard valence excitations should not be accompanied by considerable orbital relaxation. Second, the situation is different for Rydberg states, where TDDFT has long been known to systematically underestimate excitation energies on account of delocalization error [251]. ROKS/RO-ΔSCF dramatically reduces errors in local functionals, often by more than a factor of 3. The residual error still stems from systematic underestimation, which is likely on account of delocalization error (which overstabilizes the diffuse density of Rydberg states relative to the ground state), This is similar to behavior seen for CT excited states in Ref 265. The global

Method	Valence Excitations		Rydberg Excitations		All Excitations	
DFT protocols	RMSE	ME	RMSE	ME	RMSE	ME
SPW2	0.40	-0.16	0.31	-0.30	0.35	-0.24
SPW92/TDDFT	0.73	-0.43	1.26	-1.25	1.07	-0.91
PBE	0.39	-0.29	0.45	-0.45	0.43	-0.38
PBE/TDDFT	0.35	-0.31	1.42	-1.41	1.11	-0.95
B97M-V	0.24	-0.08	0.25	-0.21	0.25	-0.16
B97M-V/TDDFT	0.54	-0.25	0.87	-0.85	0.75	-0.60
SCAN	0.50	-0.34	0.35	-0.34	0.42	-0.34
SCAN/TDDFT	0.49	-0.09	0.65	-0.61	0.59	-0.39
PBE0	0.42	-0.19	0.28	-0.27	0.34	-0.23
PBE0/TDDFT	0.46	-0.29	0.56	-0.54	0.52	-0.43
ωB97X-V	0.33	-0.01	0.12	0.08	0.23	0.04
ωB97X-V/TDDFT	0.27	-0.13	0.16	-0.15	0.22	-0.14
Wavefunction Theories [440]						
CIS(D)	0.20	0.02	0.48	-0.44	0.39	-0.25
CIS(D_∞)	0.09	-0.01	0.67	-0.67	0.52	-0.39
ADC(2)	0.09	-0.01	0.67	-0.67	0.52	-0.39
CC2	0.13	0.11	0.62	-0.62	0.48	-0.32
CCSD	0.12	0.04	0.01	-0.01	0.08	0.01
ADC(3)	0.24	-0.19	0.35	0.35	0.31	0.13
CC3	0.03	0.00	0.04	-0.04	0.04	-0.02

Table 11.3: Errors (in eV) in predicting low lying excited states of HCHO (as given in Table 11.2) for various functionals, using both TDDFT (as indicated in the table) and ROKS/RO-ΔSCF. wavefunction theory errors have been found from values in Table S6 of Ref 440.

hybrid PBE0 also sees a substantial reduction in error with the orbital optimized procedure, although the range separated hybrid ωB97X-V functional gives very similar behavior across both approaches. This is not too surprising, as the non-local exchange in ωB97X-V guarantees correct asymptotic behavior for long-ranged particle-hole interactions (that are essential for Rydberg states) within linear-response theory itself.

The overall ROKS/RO-ΔSCF DFT errors for the Rydberg states compare very well with the wavefunction theory errors in Table 11.3, with only the highly expensive CCSD and CC3 methods having substantially lower errors. The wavefunction theories however are more accurate for the valence excitations, although the ROKS/RO-ΔSCF errors are not too large for some modern functionals. Further studies involving larger datasets (like the full set presented in Ref 440) and many more functionals would be necessary to determine the overall efficacy of the DFT based approaches. We note that the performance of ROKS with LSDA, PBE, B3LYP and ωB97X-V has been recently reported for singlet states of that dataset [441], with the results indicating that ROKS has slightly lower errors compared to TDDFT errors reported elsewhere [319], largely through improving Rydberg state predictions. Indeed,

ωB97X-V/ROKS is found to have similar RMSE [441] as wavefunction methods like ADC(2), CC2 or CIS(D). The relatively low errors of ROKS/RO-ΔSCF and the high computational scaling of wavefunction based methods however indicate considerable promise for use of ROKS/RO-ΔSCF to study low-lying excited states of large systems where wavefunction theory is unaffordable and TDDFT unsuitable.

11.5 Low Energy Excited States of Zinc Phthalocyanine

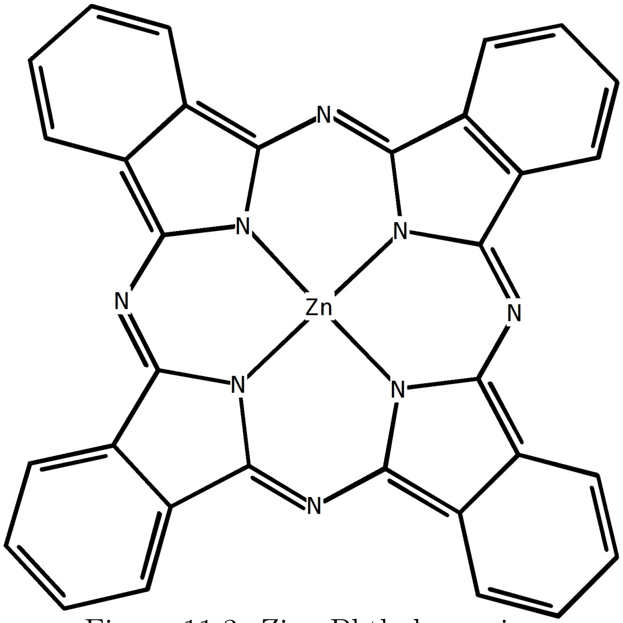


Figure 11.3: Zinc Phthalocyanine.

Metallophthalocyanines are species with a large, extensively π conjugated phthalocyanine ligand coordinated to a metal atom. They share many structural features with the biologically relevant porphyrin species and possess readily tunable electronic properties that has led to use in the electronics industry [442–444] and as photosensitizers [445]. The excited state spectra of Zn-phthalocyanine (ZnPc) has consequently been studied both experimentally [446–450] and theoretically [451–455]. The sheer size of this system makes DFT based approaches the reasonable choice (although a coupled cluster study with very small basis and significant virtual space truncation has been reported [455]) and we have consequently chosen it to demonstrate SGM's applicability to sizeable systems. We examined the symmetry allowed singlet excited states involving the twenty highest energy occupied orbitals and the ten lowest energy virtual orbitals, using the PBE0 functional and the def2-SV(P) [438] basis. The corresponding TDDFT excitation energies were also computed for comparison. Only the low energy (i.e. $\leq$ 5 eV) Q,B and N bands are reported here [446], as high energy states have (potential) multiconfigurational character and the possibility of Rydberg like behavior [455] that cannot be captured without diffuse functions in the basis set. It is worth noting that ligand to metal CT (LMCT) type transitions are not possible as Zn has a full $3d^{10}$ shell. Metal to ligand CT (MLCT) transitions are possible, but appear to occur at energies ≥ 5.3 eV with both ROKS and TDDFT. The reported transitions are therefore entirely based out of ligand orbitals.

Transition	ROKS	TDDFT	Experiment [450]	
$2a_{1u}$(HOMO) $\rightarrow 7e_g$(LUMO)	1.97	2.15	1.89	Q
			2.08	Q′
$3b_{2u} \rightarrow 7e_g$(LUMO)	3.57	3.56		
$6a_{2u} \rightarrow 7e_g$(LUMO)	3.72	3.86	3.71	B_2
$2b_{1u} \rightarrow 7e_g$(LUMO)	3.96	3.92		
$5a_{2u} \rightarrow 7e_g$(LUMO)	3.97	4.03	3.74	B_1
$28e_u \rightarrow 7e_g$(LUMO)	4.08	3.99	3.99	B_3
$2a_{1u}$(HOMO) $\rightarrow 8e_g$	4.13	4.10		
$1a_{1u} \rightarrow 7e_g$(LUMO)	4.52	4.52	4.41	N_1
			4.7	N_2

Table 11.4: Symmetry allowed vertical singlet excitation energies (in eV) computed with PBE0/def2-SV(P), as compared to experimental measurements of ZnPc in Ar matrix [450]. The experimental band assignments (Q,B etc.) have also been supplied. The TDDFT assignments have been based on the largest coefficient for transitions, in the ground state orbital basis.

The computed vertical excitation energies have been reported in Table 11.4, along with Ar matrix experimental data [450]. The computed ROKS and TDDFT energies agree very well with each other, showing that orbital optimization was not particularly necessary for this system. Nonetheless, the good agreement between the two approaches permits us to draw conclusions more confidently, as TDDFT possesses multiconfigurational character (within the singles subspace). The lack such multiconfigurational character therefore does not appear to affect ROKS performance here.

The lowest energy Q band for ZnPc is well separated from the rest of the spectrum, on account of the HOMO and LUMO being energetically separated from other orbitals. Our computed PBE0/def2-SV(P) ROKS energy for the Q band agrees quite well with experiment. However, we do not observe any symmetry allowed states that are close in energy to the Q′ state that has been suggested by experimental work (via subtraction of simulated curves from observed spectra) [450]. Previous theoretical work [451, 452, 455] has also not observed such a state, indicating that it is not of electronic origin. There exists a possibility that it is a symmetry forbidden state that appears due to vibronic distortion. However, the original assignment of it being of a A_{2u} state stemming from excitation of N lone pairs into π^* levels [449, 450] is very unlikely on account of lack of any lone pairs close in energy to the HOMO (even after ignoring symmetry considerations). It is also worth noting that the HOMO^2 →LUMO^2 double excitation energy is estimated to be 3.56 eV by ΔSCF with PBE0/def2-SV(P), making it an unlikely candidate for the Q′ band. This dark state can nonetheless play a role in the photophysics/photochemistry of the system. It is, of course inaccessible to TDDFT, which illustrates a comparative strength of the ΔSCF approach.

The B band is experimentally observed to be very broad, extending from approx. 3.0 eV to 4.3 eV. Ref 450 interpreted it as a combination of two transitions B_1 and B_2, although solvent phase measurements have suggested the presence of as many as five separate transitions [449]. We also find 5 states with E_u symmetry corresponding to that region of the spectrum, with energies roughly centered around the reported Ar matrix band maximums. Interestingly, one of those states is a transition to an unoccupied orbital that is not the LUMO (the $2b_{1u} \rightarrow 8e_g$ excitation). It is worth noting that the B_3 state is distinct from the rest of the B band as it has been assigned to be a N lone pair to π^* transition of A_{2u} symmetry, and we also find a state with the same symmetry at 4.08 eV with ROKS, offering fairly reasonable agreement.

The N band offers more of a challenge, for although we observe a state similar to the experimentally observed N_1 state, no state anywhere close in energy to the N_2 band was found with either TDDFT or ROKS. This might be a consequence of the multiconfigurational nature of the state (which could cause the ROKS energy to be too high). However, the N_2 band was a very weak contributor to the experimentally observed N band, and could likely have a non-electronic origin (or arise from symmetry forbidden transitions on account of vibronic perturbations).

11.6 Summary and Discussion

We have presented a general approach to converge excited state solutions for *any* quantum chemistry orbital optimization technique. A simple finite difference based implementation of the resulting Squared Gradient Minimization (SGM) approach requires only analytic orbital gradients of the energy/Lagrangian and costs approximately three times as much as standard ground state minimization (on a per iteration basis). SGM represents a direct minimization based alternative to the existing Maximum Overlap Method (MOM) [341, 344], that provides robust minimization to the stationary point closest to the initial guess at the expense of somewhat increased computational cost. It is simpler and thus more efficient, though also more initial-guess dependent, than other recently proposed excited state variational principles. [352, 353, 369]

Promising results were obtained within the KS-DFT framework (using the ΔSCF and ROKS approaches), especially for challenging problems like charge-transfer, Rydberg and doubly excited states (using ΔSCF when no electron pairs are broken and ROKS when one pair is uncoupled) that are beyond the ability of standard TDDFT to model. TDDFT nonetheless possesses the distinct advantage of being ‘black-box’ in the sense that it permits simultaneous computation of multiple excited states without any prior knowledge about their nature/energies. TDDFT is also quite accurate for low lying valence excitations of closed-shell molecules, where state-by-state orbital optimization offers little additional benefit. It is therefore useful to list the circumstances under which usage of SGM is likely to be beneficial for applications purposes.

SGM is the most effective when the nature of the target state can be reliably guessed, from chemical intuition or experimental data. The Q band of ZnPc is a clear example of this nature, as it is quite well understood to be a HOMO→LUMO type of transition. Similarly, it is also often possible to enumerate potential CT states in donor-acceptor complexes or LMCT/MLCT excitations in transition metal compounds, and directly target them. Naive enumeration of states would likely be unwise on account of a rapidly growing number of possibilities, necessitating use of narrow selection criteria to truncate the search space.

An alternative is to first run a pilot TDDFT computation and subsequently determine which states are of CT or Rydberg nature, followed by specifically optimizing them with SGM while leaving valence excitations as is. TDDFT natural transition orbitals (NTOs) could in fact prove to be very useful initial guesses for such problems. This strategy would not be useful for double excitations (as TDDFT cannot detect them directly). The best way to identify potential double excitations (aside from chemical intuition) is via examination of low energy TDDFT single excitations, which might couple together.

Further work is certainly desirable to assess the performance of ground state functionals for modeling excited states within the ΔSCF and ROKS approaches. It is possible to consider extending the present approach by employing SGM to converge regularized orbital optimized MP2 (OOMP2) [276] or even orbital optimized coupled cluster (CC) methods [419] for excited states. This direction potentially complements very recent work on converging CC amplitudes for excited states. [456, 457]

From a practical standpoint, the ΔSCF and ROKS ansatze constrain the current applications of SGM to single configuration excited states. This limitation can be lifted in practice by using a set of optimized excited HF determinants (i.e. abandoning DFT) as a many-electron basis for non-orthogonal Configuration Interaction (NOCI) [394, 395]. NOCI-MP2 [458, 459] then provides an approach to add dynamic correlation in a well defined manner and for relatively low computational cost.

11.7 Computational Details

All calculations were performed with the Q-Chem 5.2 [191] package. Local exchange-correlation integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points for all atoms.

11.8 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: geometries of species studied, TDDFT excitation energies and provenance of geometries.

Chapter 12

Restricted Open-Shell Kohn-Sham for Core-Level Excitations

12.1 Main Text

Spectroscopy of core electrons is an useful tool for characterizing local electronic structure in molecules and extended materials, and has consequently seen wide use for studying both static properties [370–372] and dynamics [373–375] of chemical systems. Theoretical modeling of core excited states is however a challenging task, as traditional quantum chemistry methods are typically geared towards understanding behavior of valence electrons. Indeed, it is common practice to 'shift' computed X-ray absorption spectra (XAS) by several eV to align with experiment [328, 460–464]. Such uncontrolled translation of spectra for empirical mitigation of systematic error is quite unappealing, and creates considerable scope for incorrect assignments.

Linear response (LR) methods like time-dependent density functional theory (TDDFT) [250, 251, 306] and equation of motion coupled cluster (EOM-CC) [381, 382] are widely used to model excitations. LR methods do not require prior knowledge about the nature of targeted states, as they permit simultaneous calculation of multiple states on an even footing. However, widely used LR methods only contain a limited description of orbital relaxation, leading to poor performance for cases where such effects are essential (such as double excitations [325, 326, 355], as well as charge-transfer [251, 387] and Rydberg states [325, 406] in the case of TDDFT). Core excitations in particular are accompanied by substantial relaxation of the resulting core-hole (as well as relaxation of the valence density in response), leading to rather large errors with standard LR protocols. For instance, TDDFT spectra often need to be blue-shifted by > 10 eV to correspond to experiment [328, 460–462] (unless short-range corrected functionals specifically trained to predict core spectra are employed [377]) and even EOM-CC singles and doubles (EOM-CCSD) [381] tends to systematically overestimate energies by $1 - 2$ eV [378, 465]. It worth noting that second order algebraic diagrammatic construction (ADC(2) [466], specifically CVS-ADC(2)-x [460]) has been able to attain better accuracy for core-excitations, but only via compensation of basis set incompleteness errors with lack of orbital relaxation [460]. The $O(N^5)$ computational scaling of ADC(2) also restricts applicability to large systems, relative to the lower scaling of DFT. Methods based on the GW approximation [467] and the Bethe-Saltpeter equation (BSE) [468] have also been employed to study core spectra [469–471].

Orbital optimized (OO) methods attempt to incorporate the full effect of orbital relaxation on target excited states. The state specificity of OO methods necessitate prior knowledge about the nature of targeted states, making them not truly black-box. They have also historically been prone to 'variational collapse' down to the ground state (as excited states are usually optimization saddle points), though recent advances in excited state orbital optimization have mitigated this to a great extent [341, 345, 352, 353]. OO methods have nonetheless been employed successfully for core ionizations [456, 457, 472, 473] and core excitations [416, 474, 475]. There also exist LR methods that incorporate partial OO character, like Static Exchange (STEX) [476] or Non-orthogonal Configuration Interaction Singles (NOCIS) [376, 477], though such treatments are wavefunction based and ~ 1 eV error remains common due to lack of dynamic correlation. Accurate single-point energies obtained from OO methods can also be employed to non-empirically translate LR excitation spectra into better agreement with experiment [478, 479].

The most widely used OO approach for modeling core excitations is Δ Self-Consistent Field (ΔSCF) [264, 341, 358, 416], where a non-aufbau solution to the Hartree-Fock [14] or Kohn-Sham [48] DFT equations is converged. Unfortunately, single excitations in closed shell molecules cannot be represented by a single Slater determinant, resulting in spin-contaminated "mixed" ΔSCF solutions that are intermediate between singlet and triplet. The core-hole and the excited electron are nonetheless expected to be fairly independent (due to low spatial overlap between orbitals), and spin-contaminated ΔSCF solutions can therefore be reasonably purified to a singlet via approximate spin-projection (AP) [357]. AP however entails independent optimization of the triplet state, resulting in two sets of orbitals per targeted singlet state, which is both computationally inefficient and intellectually unappealing. Furthermore, spin-unrestricted DFT can exhibit rather unusual catastrophic failures with electronic configurations with two unpaired electrons [218], making a restricted approach preferable.

Restricted Open-Shell Kohn-Sham [362, 363](ROKS) solves both of these issues via optimizing $2E_M - E_T$ for the same set of spin-restricted (RO) orbitals (E_M is the energy of the mixed Slater determinant and E_T is the energy of the corresponding triplet determinant within the $M_s = 1$ manifold). Most ROKS implementations (such as the one described in Ref 363) however tend to collapse down to the lowest excited singlet (S_1) state, hindering use for studying core excitations. The recently developed Square Gradient Minimization (SGM) approach [345] permits ROKS to target arbitrary singlet excited states with one broken electron pair, thereby making application to core excitations feasible. SGM has been described in detail elsewhere [345], and we only note that each SGM iteration has a cost that ranges between twice (for methods with analytical orbital Hessians for the energy/Lagrangian) and thrice (for methods without such Hessians, necessitating use of a finite-difference approximation) the cost of evaluating the orbital gradient of the energy/Lagrangian. ROKS calculations with SGM therefore have the same scaling as ground state DFT calculations with methods like GDM [348] or DIIS [342], but with a slightly larger prefactor per iteration. SGM is also robust against variational collapse and can converge to excited states where the more conventional Maximum Overlap Method (MOM) [341] encounters variational collapse or fails to converge [345].

A rather important consideration for use of ROKS is the choice of a functional out of the vast DFT alphabet soup. This is especially relevant for core spectroscopy, as modern DFT functionals have been trained/assessed mostly on modeling ground state energetics [98, 111, 112] and properties [80, 95, 226], which only depend on behavior of *valence* electrons. It therefore seems appropriate to consider non-empirical density functionals like LSDA [149], PBE [57] and PBE0 [91], or minimally parametrized functionals like SCAN [62] or ωB97X-V [108] that are fairly strongly constrained within functional space. It also seems worthwhile to assess the performance of highly accurate modern functionals like B97M-V [63], that are less tightly constrained. We have consequently examined the performance of these six functionals in predicting ROKS excitation energies for 40 K edge transitions (i.e. from the 1s orbital) of C,N,O and F, for which relativistic effects are expected to be small. The resulting values have been listed in in Table 12.1, while statistical measures of error have been provided in Table 12.2. Table 12.2 also lists errors in core ionization potentials (core IPs) and term values (gap between K edge and core IP), in order to give a more complete idea about the full spectrum. This indirect measure is useful, since it is typically difficult to identify individual transitions beyond the edge from experimental spectra. We do however note that ROKS/SGM can converge to higher excited states beyond the K edge with ease, preserving similar levels of accuracy (examples provided in Supporting Information).

The values in Table 12.2 make it quite clear that the SCAN and ωB97X-V functionals are highly accurate in predicting the K edge, having an RMSE on the order 0.3 eV irrespective of the presence of atom specific relativistic shifts. ωB97X-V appears to be a bit less accurate for the prediction of core IPs than SCAN, but the greater variation in experimental measurements of core IPs [494] indicates that not too much meaning should be drawn from this. The classic PBE0 functional also appears to be fairly accurate when relativistic effects are included (although the K edge RMSE is about twice as large as that of ωB97X-V). The SPW92 [138, 149] LSDA functional systematically underestimates energies by > 4 eV, on account of it only being exact for the uniform electron gas and therefore incapable of modeling the inhomogeneities present in the densities of core excited states. The PBE generalized gradient approximation (GGA) systematically underestimates energies by about an eV, while the B97M-V meta-GGA surprisingly appears to systematically *overestimate* by > 1.5 eV. Finally, all functionals predict term values to approximately the same accuracy, indicating that empirically translating ROKS spectra by functional specific constant shifts would lead to similar levels of accuracy, irrespective of the functional used. We however feel that translation of spectra is rather unappealing and will not pursue that avenue further.

The high accuracy predicted by SCAN and ωB97X-V (relative to experimental errors, which are on the order of 0.1 eV) merits further analysis to determine the factors responsible,

Species	Core orbital	Expt.	SPW92	PBE	B97M-V	SCAN	PBE0	ωB97X-V
C_2H_4	C	284.7 [480]	281.1	284.0	286.4	284.7	284.3	285.1
HCHO	C	285.6 [481]	282.1	284.9	287.4	285.7	285.2	286.0
C_2H_2	C	285.9 [480]	282.1	284.8	287.3	285.7	285.2	286.0
C_2N_2	C	286.3 [482]	282.5	285.3	287.8	286.2	285.7	286.6
HCN	C	286.4 [482]	282.8	285.5	288.0	286.3	285.8	286.6
$(CH_3)_2CO$	C (CO)	286.4 [483]	282.9	285.6	288.1	286.4	285.9	286.6
C_2H_6	C	286.9 [480]	282.8	285.8	288.1	286.7	286.3	287.3
CO	C	287.4 [484]	283.5	286.1	288.7	287.0	286.5	287.3
CH_4	C	288.0 [485]	284.0	286.9	289.4	288.0	287.4	288.5
CH_3OH	C	288.0 [483]	284.6	287.5	289.9	288.2	287.7	288.7
HCOOH	C	288.1 [483]	284.2	287.0	289.6	288.0	287.4	288.2
HCOF	C	288.2 [486]	284.4	287.2	289.8	288.1	287.6	288.4
CO_2	C	290.8 [487]	286.5	289.1	292.0	290.3	289.7	290.5
CF_2O	C	290.9 [486]	286.8	289.5	292.3	290.6	290.0	290.8
C_2N_2	N	398.9 [482]	394.5	397.8	400.5	398.7	398.2	399.1
HCN	N	399.7 [482]	395.4	398.7	401.3	399.5	399.0	399.8
Imidazole	N (CH=N-CH)	399.9 [488]	395.6	398.9	401.5	399.7	399.2	399.9
NH_3	N	400.8 [485]	395.9	399.4	402.0	400.3	399.8	400.9
N_2	N	400.9 [489]	396.6	399.8	402.5	400.7	400.1	400.9
N_2O	N ($\overset{*}{N}NO$)	401.0 [487]	396.7	400.0	402.7	400.9	400.2	401.0
Glycine	N	401.2 [490]	396.5	400.0	402.6	400.9	400.5	401.6
Pyrrole	N	402.3 [491]	397.8	401.3	403.9	402.2	401.7	402.5
Imidazole	N (CH-NH-CH)	402.3 [488]	397.9	401.3	403.9	402.2	401.7	402.5
N_2O	N ($N\overset{*}{N}O$)	404.6 [487]	400.0	403.3	406.1	404.4	403.7	404.5
HCHO	O	530.8 [481]	525.9	529.8	532.5	530.6	529.9	530.8
$(CH_3)_2CO$	O	531.4 [483]	526.2	530.1	532.8	531.0	530.3	531.1
HCOF	O	532.1 [486]	527.0	530.9	533.6	531.8	531.0	531.9
HCOOH	O(CO)	532.2 [483]	526.9	530.8	533.5	531.7	530.9	531.8
CF_2O	O	532.7 [486]	527.9	531.9	534.7	532.8	532.0	532.9
H_2O	O	534.0 [485]	528.6	532.5	535.4	533.6	533.0	534.0
CH_3OH	O	534.1 [483]	528.8	532.7	535.5	533.8	533.2	534.1
CO	O	534.2 [484]	529.1	533.0	535.7	533.9	533.1	534.0
N_2O	O	534.6 [487]	529.9	533.9	536.7	534.8	533.9	534.6
Furan	O	535.2 [492]	530.0	534.0	536.6	534.9	534.2	535.1
HCOOH	O(OH)	535.4 [483]	530.1	534.2	537.0	535.2	534.5	535.4
CO_2	O	535.4 [487]	530.3	534.2	537.1	535.3	534.4	535.3
F_2	F	682.2 [493]	676.8	681.2	683.9	682.0	681.1	682.0
HF	F	687.4 [493]	681.4	685.8	688.9	687.1	686.2	687.1
HCOF	F	687.7 [486]	681.8	686.3	689.3	687.5	686.5	687.5
CF_2O	F	689.2 [486]	683.4	687.9	691.0	689.1	688.1	689.1

Table 12.1: Comparison between experimental (Expt.) and ROKS/aug-cc-pCVTZ K edge (lowest symmetry allowed transition from 1s core orbitals) vertical absorption energies, of 40 core excitations in small molecules (in eV).

Functional	K edge		K edge (+rel. corr.)		Core IP (+rel. corr.)		Term value	
	RMSE	ME	RMSE	ME	RMSE	ME	RMSE	ME
SPW92	4.6	-4.6	4.4	-4.3	4.2	-4.2	0.3	0.2
PBE	1.2	-1.1	0.9	-0.9	0.8	-0.8	0.3	0.1
B97M-V	1.6	1.6	1.8	1.8	1.8	1.8	0.3	0.1
SCAN	0.2	-0.2	0.2	0.1	0.3	0.2	0.3	0.2
PBE0	0.9	-0.8	0.6	-0.6	0.4	-0.4	0.4	0.2
ωB97X-V	0.2	0.1	0.4	0.3	0.5	0.4	0.4	0.1

Table 12.2: Root mean squared error (RMSE) and mean error (ME) for prediction of K edge energies listed in Table 12.1 (in eV). The effect of relativistic corrections (rel. corr.) have also been considered. The errors in prediction of the corresponding core ionization potential (core IP) and the term value (difference between K edge and core IP) are also reported.

and what error cancellations (if any) are occurring. Some of the most obvious factors to consider are relativistic effects, the roles played by orbital relaxation and delocalization error, as well as basis set incompleteness errors. Scalar relativistic effects systematically bind core electrons tighter than what predictions from non-relativistic DFT should suggest. The magnitude of this correction can be estimated from the difference between core IPs calculated with relativistic and non-relativistic theories for bare atoms. This approximation should be fairly accurate for second period elements, as the chemical environment would only slightly perturb these already small corrections (the reported values [495] range from 0.1 eV for C to 0.6 eV for F). Inclusion of these relativistic shifts however has minimal impact on the K edge RMSE for SCAN and ωB97X-V (as can be seen from Table 12.2), as well as for core IPs (as shown in the Supporting Information). The corrections do however appear to perceptibly lower RMSE for PBE0, by reducing some of the systematic underestimation. We also note that relativistic corrections are expected to be much larger past the second period, and cannot be neglected in K edge studies of heavier atoms.

The overall effect of explicit orbital optimization via ROKS can be gauged by comparison to LR-TDDFT. Table 12.3 presents the results for the CH_4 and HF molecules, which conclusively demonstrate the utility of orbital optimization (as TDDFT underestimates experiment by 15-20 eV). We also note that ΔSCF has similar accuracy as ROKS, showing that the coupling between the core-hole and excited electron is indeed very weak. Our conclusions about the behavior of ROKS with various functionals are therefore likely transferable to ΔSCF in the regimes where the latter does not exhibit any unphysical behavior.

The poor performance of TDDFT naturally raises questions about the role of delocalization error [64] (of which self-interaction error is but one part [83, 95]), which is the factor typically responsible for systematic underestimation of TDDFT excitation energies [251]. The excellent behavior of the SCAN meta-GGA local functional, and the relatively small performance gap between the local PBE and the global hybrid PBE0 functionals seem to

	HF	CH_4
Experiment	687.4	288.0
SCAN/TDDFT	666.1	273.8
SCAN/ΔSCF	687.1	287.9
SCAN/ROKS	687.0	288.0
ωB97X-V/TDDFT	668.7	276.5
ωB97X-V/ΔSCF	687.2	288.5
ωB97X-V/ROKS	687.1	288.5

Table 12.3: Comparison of TDDFT, ΔSCF and ROKS K edges (in eV) for HF and CH_4 with the SCAN/ωB97X-V functionals and the aug-cc-pCVTZ basis. The ΔSCF values have been spin-purified with AP.

suggest that delocalization error is not a major factor for ROKS. This is consistent with earlier observations of ROKS predicting excellent charge-transfer [265] and Rydberg [345] state energies for cases where TDDFT performs poorly. Delocalization error of course continues to exist for ROKS, but orbital optimization drastically reduces the magnitude of delocalization driven errors that LR methods tend to predict [265, 321, 407], down to ground state calculation levels.

	Delocalized hole	Localized hole	Difference
SPW2	388.4	396.6	-8.2
PBE	391.7	399.8	-8.1
B97M-V	398.9	402.5	-3.6
SCAN	395.3	400.7	-5.4
PBE0	396.4	400.1	-3.7
ωB97X-V	395.7	400.9	-5.3
Experiment		401.0	

Table 12.4: Comparison of the N_2 K edge predicted by ROKS/aug-cc-pCVTZ (in eV) between the fully delocalized and fully localized core-hole limits.

There is however an additional subtlety associated with systems possessing chemically identical atoms (like N_2 or O in CO_2), where the core-hole density of the exact eigenstate should be delocalized over multiple sites on account of symmetry. The coupling between core orbitals is nonetheless quite weak and localized core-hole diabatic states are therefore expected to be energetically quite close (i.e. within order of 0.01 eV) [496] to symmetric eigenstates. The energies of delocalized states in DFT are typically systematically underestimated on account of delocalization error (even within an OO framework), making use of localized core-hole states preferable for calculating core excitation energies. A quantitative measure of this effect for the N_2 molecule has been supplied in Table 12.4. In practice therefore, the spurious delocalization effect should be avoided by supplying a localized core-hole as the initial guess and letting SGM converge to the closest localized solution. *However*, it means that canonical orbitals cannot be used as initial guesses due to their inherently delocalized nature, and some localization scheme (or even a weak, symmetry breaking electric field) must be employed to generate initial guess orbitals for ROKS. It is somewhat intellectually unsatisfying to completely neglect delocalized states (which appear to be the lowest energy ROKS core-hole states as well as representative of the behavior expected from the true eigenstate), but this pragmatic choice is essential in light of known failures of DFT for delocalized states [78–81]. Fully symmetric states can be obtained from a NOCI approach [376, 477, 496], but such multireference techniques cannot be straightforwardly generalized to DFT. We additionally note that localized orbitals has long been employed to improve the performance of wavefunction based approaches as well [376, 497], although use of delocalized orbitals therein lead to *higher* energies (on account of missing correlation [95]). The actual energy gap between the exact eigenstate with a delocalized core-hole and a localized core-hole state is however quite small overall [496], and therefore use of localized ROKS solutions is an acceptable pragmatic choice. We additionally note that this small gap indicates that any experimental realization of a localized core-hole state (due to finite-temperature effects or other symmetry breaking) in experiment would not affect accuracy of experimental data employed in this study.

	aug-cc-pCVDZ		aug-cc-pCVTZ		aug-cc-pCVQZ	
	Core IP	K edge	Core IP	K edge	Core IP	K edge
CH_4 (C)	292.07	289.42	291.16	288.50	291.11	288.44
NH_3 (N)	407.02	402.01	405.83	400.85	405.76	400.78
H_2O (O)	541.33	535.44	539.86	533.99	539.75	533.88
HF (F)	695.63	688.86	693.87	687.13	693.72	686.98
HCHO (C)	295.89	286.97	294.97	286.03	294.91	285.97
HCHO (O)	540.76	532.32	539.27	530.83	539.15	530.71
HCN (C)	295.03	287.71	293.9	286.58	293.84	286.51
HCN (N)	408.37	401.09	407.07	399.8	406.99	399.71

Table 12.5: Convergence of ωB97X-V core ionization potential (IP) and K edge absorption energies (in eV) against basis set size.

The final factor we consider is basis set incompleteness error, whose analysis would also assist basis set selection for realistically sized systems (as aug-cc-pCVTZ is too impractically large). Valence excitation energies typically do not exhibit very strong basis set dependence [440], but the situation for core spectra is different due to the need to adequately relax the core-hole. Table 12.5 compares the ωB97X-V core ionization and K edge energies with increasing basis set cardinality. The small difference between aug-cc-pCVTZ and aug-cc-pCVQZ and the exponential convergence of SCF energies [239] suggest that aug-cc-pCVQZ values are functionally at the complete basis set (CBS) limit. It can also be seen that aug-cc-pCVTZ systematically overestimates energies by about 0.1 eV relative to aug-cc-pCVQZ.

This deviation is non-negligible relative to the low RMSE of SCAN and ωB97X-V, but is quite comparable to the error bars inherent in experiment, indicating that the basis set incompleteness error in Table 12.1 is not particularly significant. We nonetheless note that the slight overestimation of energies by aug-cc-pCVTZ seems to suggest that a component of the systematic overestimation of energies (after relativistic corrections) for SCAN and ωB97X-V stems from basis set truncation, suggesting slightly lower errors at the CBS limit.

Table 12.5 also makes it apparent that aug-cc-pCVDZ is too small for accurate predictions, as energies are systematically overestimated by 1-2 eV. The core IP is overestimated by almost the same amount as the K edge, indicating that the basis set incompleteness effects essentially arise from insufficient core relaxation alone. We therefore recommend that a mixed basis strategy be employed for larger species (where full aug-cc-pCVTZ is impractical), wherein the localized target atom employs a split core-valence triple zeta (CVTZ) basis, while the remaining atoms are treated with some smaller basis. A similar mixed basis approach has also been reported in literature [498]. This strategy (using aug-cc-pCVTZ in the target site and aug-cc-pVDZ for other atoms) reproduced the full ωB97X-V/aug-cc-pCVTZ results for both the C and O K edges of HCHO to ≤ 0.02 eV deviation, suggesting its general efficacy for predicting K edges of second period elements. It is however important to recognize that CVnZ bases are not likely to be sufficiently flexible in describing 1s electrons beyond the second period, and more specialized (or even uncontracted) basis sets may prove necessary for the local site of the K shell excitation for heavy elements. In addition, we note that while diffuse functions are not strictly necessary for excitations to antibonding orbitals, they are critical for Rydberg states, with double augmentation being necessary to converge the higher core excited states of small molecules (as shown in the Supporting Information).

We next demonstrate the viability of applying ROKS/SGM to sizeable systems by computing the N K edge of the phthalocyanine molecule (H_2Pc, depicted in Fig 12.1). We employ the mixed basis strategy described and validated earlier, with the large cc-pCVTZ basis being applied to the target site while all other atoms use cc-pVDZ. We note that an additional advantage of the mixed basis approach is that it automatically breaks chemical equivalence of the target site, thereby spontaneously localizing the resulting core orbital (sans explicit localization). Fig 12.1 shows that H_2Pc has three different types of N atoms. N1-N4 are bridging aza type, N7-N8 are NH pyrrole like while N5-N6 are hydrogen free pyrrole like. A comparison between ωB97X-V/ROKS excitation energies and experimental values from thin film measurements [499] are supplied in Table 12.6. We continue to find remarkably good agreement between theory and experiment, with the N core energies being predicted to be in the order N5<N1<N7.

Having discussed the applicability of using ROKS/SGM for the computation of 1s excitation energies for second period elements, we briefly consider the behavior of inner shell excitations for heavier atoms. Excitations out of the 2p orbitals are of particular interest for third period elements. The degeneracy of the 2p orbitals is however broken by spin-orbit coupling (which is larger in magnitude than any splitting introduced by molecular symmetry on core levels), which results in two peaks with intensities roughly in a 2:1 ratio, that correspond to excitations out of the $2p_{3/2}$ and $2p_{1/2}$ levels respectively. These peaks are called

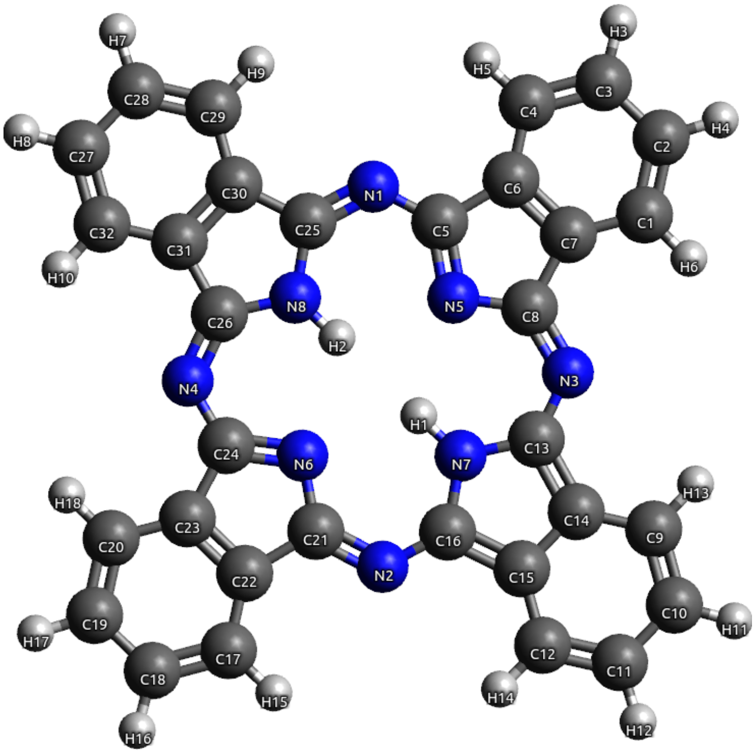

H7
H9
H5
H3
C28
C29
C4
C3
H4
H8
C27
C2
C30
N1
C6
C32
C31
C25
C5
C7
C1
H10
C26
N8
N5
C8
H6
H2
N4
N3
H1
C24
N6
N7
C13
H18
H13
C23
C21
C16
C14
C20
N2
C9
C22
C15
C10
H17
C19
H11
C18
C17
C12
C11
H15
H14
H16
H12

Figure 12.1: H_2Pc molecule.

ROKS	core-hole site	Experiment
398.3	N5	397.9
398.4	N5	398.3
398.5	N1	
399.1	N1	399
400.3	N7	399.7
400.5	N7	400.3

Table 12.6: Comparison of experimental N 1s excitation energies [499] (in eV) of H_2Pc to predictions from ROKS with ωB97X-V. A mixed basis set (see text) was used.

L_3 and L_2 respectively (in contrast to the higher energy L_1 peaks stemming from excitations out of the 2s level). This spin-orbit splitting cannot be reproduced by any non-relativistic theory like Kohn-Sham DFT. Like the scalar relativistic shifts employed earlier however, they are not sensitive to the chemical environment of a given atom. The spin-orbit effects of the electron excited to a higher energy orbital is also typically negligible on account of greater distance from the nucleus. It is therefore possible to estimate the L_3-L_2 splitting for

	L_3			L_2		
	Expt.	SCAN	ωB97X-V	Expt.	SCAN	ωB97X-V
SiH_4	102.6 [500]	103.0	102.9	103.2 [500]	103.6	103.5
$Si(CH_3)_4$	102.9 [501]	102.8	102.8	103.5 [501]	103.4	103.4
SiF_4	106.1 [502]	106.2	106.1	106.7 [502]	106.8	106.7
$\overset{*}{\mathrm{Si}}(\mathrm{Cl})_4$	104.2 [501]	104.5	104.6	104.8 [501]	105.1	105.2
$Si(OCH_3)_4$	104.8 [503]	104.9	105.1	105.4 [503]	105.5	105.7
PH_3	131.9 [504]	132.1	131.8	132.8 [504]	132.9	132.6
PF_3	134.9 [505]	134.9	134.7	135.6 [505]	135.7	135.5
$P(CH_3)_3$	132.3 [504]	132.5	132.2	133.1 [504]	133.3	133.0
PF_5	138.2 [506]	138.0	138.0	139.0 [506]	138.8	138.8
OPF_3	137.1 [505]	137.0	136.9	137.8 [505]	137.8	137.7
H_2S	164.4 [507]	164.7	164.3	165.6 [507]	165.9	165.5
SF_6	172.3 [508]	172.0	171.9	173.4 [508]	173.2	173.1
$(CH_3S)_2$	164.1 [375]	164.0	163.6	165.4 [375]	165.2	164.8
CS_2	163.3 [509]	163.4	162.5	164.4 [509]	164.6	163.7
CSO	164.3 [510]	164.4	163.7	165.5 [510]	165.7	164.9
HCl	200.9 [511]	201.0	200.5	202.4 [512]	202.6	202.1
Cl_2	198.2 [513]	198.2	197.7	199.8 [513]	199.9	199.3
ClF_3	201.8 [514]	201.7	201.3	203.2 [514]	203.3	203.0
CCl_4	200.3 [515]	200.1	199.7	201.9 [515]	201.7	201.3
C_6H_5Cl	201.5 [516]	201.4	201.0	203.2 [516]	203.1	202.6
RMSE		0.2	0.4		0.2	0.4
ME		0.1	-0.2		0.1	-0.2

Table 12.7: Comparison between experimental (Expt.) and ROKS $L_{2,3}$ edges (lowest symmetry allowed transition from 2p core orbitals) vertical absorption energies (in eV). SiH_4, PH_3, H_2S and HCl employ aug-cc-pCVTZ, while the mixed basis strategy described above was used for the remaining species (aug-cc-pCVTZ as the large local basis and aug-cc-pVDZ for other atoms). Scalar relatvistic corrections for these atoms are < 0.1 eV [495] and were thus neglected. The protocol for incorporating spin-orbit coupling is described in 12.2.

a specific atom either via relativistic wavefunction theories or experiment, and transfer those values for other species via use of (near-)degenerate perturbation theory, in conjunction with the non-relativistic values computed with ROKS (as described in 12.2). Table 12.7 supplies a comparison between values obtained with this method (employing the hereto best performing SCAN and ωB97X-V functionals) with experiment for a few species. Both functionals appear to retain the level of accuracy observed for the second period K edge. It does however appear that ωB97X-V performs a little worse than SCAN due to systematic underestimation of excitation energies. Nonetheless, it is apparent that this approach is quite promising for computing core spectra of 2p electrons in heavier elements, in addition to the second period K shell spectroscopy discussed so far.

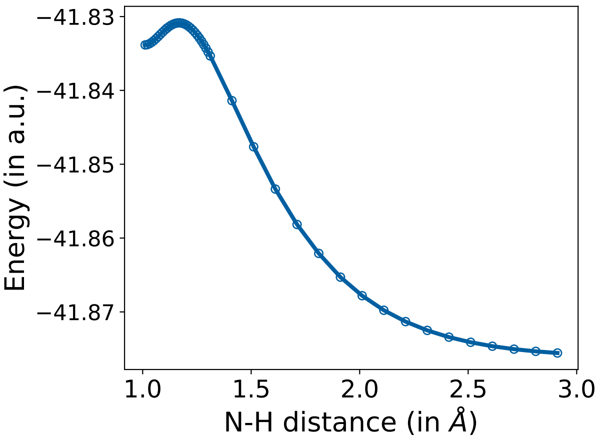


Figure 12.2: PES of core excited NH_3 (from ROKS ωB97X-V/aug-cc-pCVTZ), against stretch of a NH bond. Nuclear positions of the other atoms were optimized for all points.

It is also worth noting the analytical nuclear gradients for ROKS are fairly simple to obtain [363], permitting geometry optimizations and ab-initio molecular dynamics in the core-excited state (which could assist in studying ultrafast dissociation processes or lead to ab-initio computation of spectral linewidths, for instance). Conseuqently, it is also possible to compute vibrational spectra of core excited states via finite differences, making it possible to assign modes to vibrational fine structure of XAS. All of this can be acheived for the same computational scaling as ground state DFT, permitting application to very large systems. As a simple of proof of principle, Fig 12.2 presents the potential energy surface (PES) of core excited NH_3 (1s→ $4a_1$) against NH stretching. This state can relax to a shallow local minimum, but ultrafast dissociation to NH_2+H is energetically more favorable (after crossing a small barrier [517]). ROKS is able to reproduce this behavior, which is a significant advantage over TDDFT (as the latter is completely incapable of modeling excited state bond dissociation [228]). The barrier against dissociation is estimated to be 0.08 eV, which is within the 0.1 eV error bar associated with the experimental estimate of 0.2 eV [517]. It is however worth noting that typical DFT error for ground state barrier prediction is of the order of 0.05 eV [98], and so ultraprecise predictions should not be realistically expected. The main strength of ROKS is that it can be applied to large systems with *reasonable* accuracy.

We also demonstrate reasonable reproduction of the core excited state bond length and vibrational frequency of N_2 by ROKS, which has been fairly well characterized by both theory and experiment [489]. A comparison with the experimental values, the CC3 [439] wavefunction method (from Ref 489) and ROKS is provided in Table 12.8. We find that theoretical methods predict a shorter and stiffer bond in the core-excited state, relative to experimental fits. We do however note that the experimental values are not particularly precise, with the vibrational frequency being estimated from an experiment with a photon resolution of approx 50 meV (i.e. 403 cm^{-1}) and the bond length being calculated via a

	Expt.	CC3 [489] d-aug-cc-pCVQZ	SCAN aug-cc-pCVTZ	ωB97X-V aug-cc-pCVTZ	Ground state Expt.
Absorption energy (in eV)	400.90 [489]	401.03	400.73	400.91	
Bond length (in Å)	1.164 [518]	1.158	1.154	1.147	1.098
Frequency (in cm^{-1})	1895 [518]	2032	2049	2134	2330

Table 12.8: Comparison of CC3 and ROKS predictions for first core excited state ($1s \to \pi^*$) state of N_2.

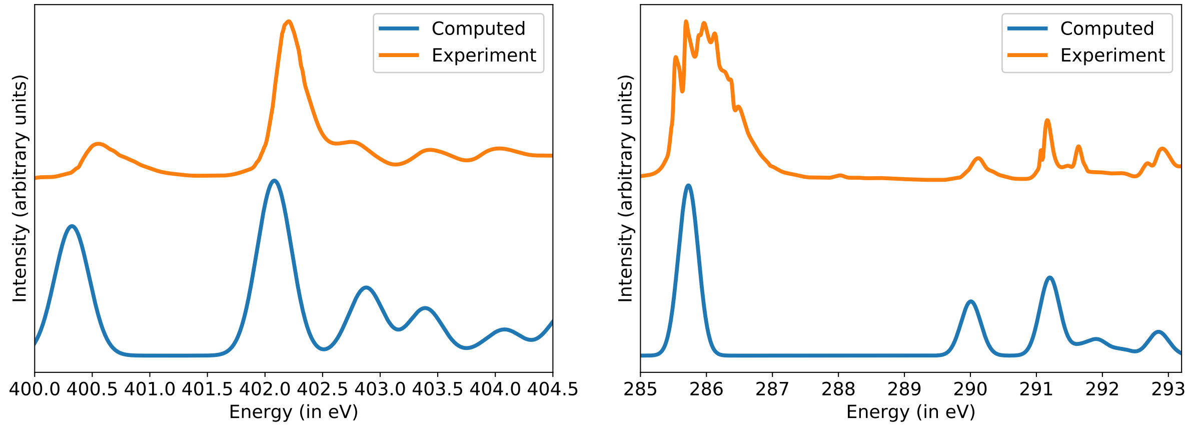
fit to a Morse potential [518], which does not appear to be consistent with coupled cluster studies [489]. The SCAN predictions are in very good agreement with CC3, while ωB97X-V predicts a shorter bond and higher vibrational frequency. This superficially seems to suggest higher reliability of SCAN geometries/frequencies, but considerable further testing is required before more general conclusions can be reached. At any rate, the low computational cost of ROKS with either functional makes it attractive relative to $O(N^7)$ scaling methods like CC3.

Molecule	Hole site	Expt.	SPW92	PBE	B97M-V	SCAN	PBE0	ωB97X-V
C_2H_2	C,C	596.0± 0.5 [519]	588.3	593.8	598.7	595.6	594.7	596.3
C_2H_4	C,C	593.3± 0.5 [519]	585.0	590.7	595.6	592.5	591.5	593.1
C_2H_6	C,C	590.0± 0.5 [519]	581.7	587.6	592.4	589.3	588.3	589.9
CO	C,O	855.4±1 [519]	846.2	852.6	858.0	854.8	853.6	855.2
CO_2	C,O	848.6± 1.2 [520]	842.1	848.6	854.3	851.1	850.0	851.6
N_2	N,N	835.9± 1 [519]	827.9	834.4	839.9	836.7	835.7	837.3
N_2O	N,N	834.2±2.1 [520]	825.1	831.6	837.4	834.1	833.4	835.2

Table 12.9: Comparison between experimental and ROKS/aug-cc-pCVTZ TSDCH core ionization energies (in eV).

It is also important to note that the ROKS is applicable to any singlet state with one broken electron pair [345], and not just the single excitations considered so far. There is unfortunately very little high quality experimental data about doubly excited core states involving second period elements. We consequently look at two site double core-hole (TSDCH) states instead, which are intrinsically open-shell (possessing one unpaired electron in each singly ionized atomic site) and are thereby ideal candidates for ROKS. TSDCH states have been long proposed as sensitive measures of chemical environment [497], leading to experimental effort towards their realization [511, 519, 520]. We present a comparison between experimental and ROKS TSDCH ionization energies in Table 12.9. Similar behavior to the K edge data in Table 12.1 is observed, with B97M-V massively overestimating, while SPW92/PBE underestimate. The large experimental error bars make it difficult to judge the relative performances of PBE0, SCAN and ωB97X-V (however, $E_{\text{PBE0}} < E_{\text{SCAN}} < E_{\omega\text{B97X-V}}$

for all species). The predictions from the latter three functionals are overall quite reliable (considering the experimental error bars), and offer an inexpensive and spin pure way to compute TSDCH excitation energies (vs, say more expensive methods like ΔCCSD(T), which does not lead to substantially enhanced accuracy for such systems [457]). This certainly represents a major advantage of ROKS over TDDFT, which is incapable of modelling doubly excited states at all [251, 326, 327].



(a) N in NH_3.

(b) C in HCHO.

Figure 12.3: Spectra predicted by SCAN/d-aug-cc-pCVTZ compared to experiment (without any translational shift). Experimental data taken from Ref 485 for NH_3 and Ref 481 for HCHO. Gaussian broadening was applied to the peaks, with $\sigma = 0.15$ eV.

Having described the accuracy of predicting energies via ROKS/SGM, we next briefly consider the challenge of predicting actual core absorption spectra. This is somewhat more of a challenge for OO based DFT methods, as transition properties like oscillator strengths cannot formally be defined within this framework (due to the fictitious nature of the Kohn-Sham determinant). Nonetheless, reasonable values can be obtained by approximating the wavefunction with the Kohn-Sham determinant, followed by computation of transition properties via a wavefunction like approach [416]. While the actual values need not be very accurate (or obey formal properties like the Thomas-Reich-Kuhn rule [251]), their relative variation is typically expected to be similar to exact behavior, resulting in roughly accurate spectral shape. Fig 12.3 presents the core excited spectra of N in NH_3 and C in HCHO against experimental results. The agreement is in no way perfect (on account of lack of nuclear quantum effects in the computed spectra, as well as use of uniform Gaussian broadening), but the main features are reproduced quite well and no translation of spectra is necessary at all. In particular, peaks corresponding to higher energy Rydberg states are quite visible, which clearly highlights SGM's ability to predict such states without variational collapse.

In summary, we demonstrate that single core excitation energies for the K shell of second period elements and $L_{2,3}$ shells of third period elements can be computed to < 0.5 eV RMS error via the use of a state specific restricted open-shell Kohn-Sham (ROKS) approach, without any need to translate spectra at all. The computational scaling of ROKS is identical to the corresponding ground state DFT calculation (with a slightly larger prefactor), when it is combined with the recently developed square gradient minimization (SGM [345]) orbital optimizer, readily permitting application to large systems. The low ROKS errors owe greatly to advances in ground state density functional development, as modern functionals like SCAN and ωB97X-V are found to be the most accurate. We further show that the low errors in prediction do not stem from basis set incompleteness errors or neglect of relativistic effects, indicating that ROKS is obtaining the right answer for the right reasons (namely that the excitation from one localized core orbital to the virtual space can be very well described by one configuration plus a description of dynamical correlation). The ready availability of analytic ROKS nuclear gradients also suggest considerable potential for employing this approach for excited state geometry optimization or ab-initio molecular dynamics. This is aided by the ability of ROKS to correctly describe excited state bond dissociations, unlike TDDFT. Finally, ROKS can be employed to double excitation or double ionization processes (where a total of one electron pair has been broken), which is difficult for LR methods.

The high accuracy and low computational scaling of ROKS makes it an ideal method for studying the dynamics of both core excited states and XAS of valence excited states in sizeable systems. ROKS (with the local SCAN functional) is also an ideal method for simulating core spectra in the condensed phase. There does however exist a need to incorporate scalar relativistic effects, in order to extend applicability to the innermost shells of heavy elements (where an additive atom specific correction might not be sensitive enough). Work along these directions is presently in progress.

12.2 Appendix: Spin-orbit effects in L-edge spectra

There are six 2p spin orbitals ($|p_{x,y,z}\rangle \otimes |\uparrow, \downarrow\rangle$ that are degenerate in non-relativistic quantum mechanics (for symmetric molecular fields and in the absence of magnetic fields). The spin-orbit coupling operator $-J\vec{L}\cdot\vec{S}$ breaks this degeneracy. It can be easily shown that the $|p_z\rangle \otimes |\uparrow\rangle$ couples with the $|p_{x,y}\rangle \otimes |\downarrow\rangle$ (and the reverse). Within this reduced subspace of 3 interacting orbitals, the spin-orbit coupling operator can be represented as:

$$\hat{H}^{(1)} = -J \begin{pmatrix} 0 & 1 & i \\ 1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \tag{12.1}$$

In most molecules however, the degeneracy between the spatial p levels is broken due to (lack of) symmetry by a small (0.1 eV scale) amount. This effect should also ideally be accounted for, and so we have a full Hamiltonian:

$$\hat{H} = \begin{pmatrix} \omega_1 & -J & -iJ \\ -J & \omega_2 & iJ \\ iJ & -iJ & \omega_3 \end{pmatrix} \tag{12.2}$$

where $\omega_{1,2,3}$ are the non-relativistic excitation energies out of the three p orbitals (as computed with ROKS or some other method). Diagonalization of this $\hat{H}$ would yield the predicted L excitation energies.

The case of $\omega_{1,2,3}$ being degenerate (i.e. symmetric molecular field) yields the well known case where the eigenvalues are $\omega - J, \omega - J$ (L_3 band) and $\omega + 2J$ (L_2 band). The 2:1 degeneracy ratio explains the standard 2:1 heights seen in experimental spectra. The separation between these ($3J$) is called the doublet splitting and is experimentally [521] found to be 0.6 eV for Si, 0.8 eV for P, 1.2 eV for S and 1.6 eV for Cl. A weak molecular field of the order of 0.1 eV therefore is unlikely to resolve separate L_3 peaks, and this was the case for species in Table 12.7. We consequently averaged the two low energy eigenvalues into a composite L_3 value for Table 12.7 (as the energy differences between those eigenvalues were < 0.1 eV). However, the full model with three separate peaks may prove necessary in some cases.

12.3 Computational Details

All calculations were performed with the Q-Chem 5.2 [191] package. Local exchange-correlation integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points. Core IPs were computed with RO-ΔSCF, which is spin-pure and equivalent to ROKS when an electron is excited to infinity. The core-ionized RO-ΔSCF orbitals were subsequently used as initial guesses for ROKS absorption energy calculations. This reduces number of ROKS iterations, by effectively decoupling the core-hole relaxation from the rest of the optimization. Such a strategy would be extremely useful for computing multiple excited states, as the core-hole relaxation process would need to be converged only once to generate initial guesses for several ROKS calculations. SGM was employed for all ΔSCF/ROKS computations. Experimental geometries (from the NIST database [192]) were used whenever possible, with geometries being optimized with MP2/cc-pVTZ in their absence (except for H_2Pc, where ωB97X-V/def2-SV(P) was used instead). Vibrational frequencies ω in Table 12.8 were found by solving the nuclear wave equation for the PES, and subsequent fitting to the anharmonic oscillator energy $E_\nu = \hbar\omega\left(\nu + \frac{1}{2}\right) - \hbar\omega x_e\left(\nu + \frac{1}{2}\right)^2$.

12.4 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: geometries of species studied, raw data.

Chapter 13

Core-Level Spectra of Open-Shell Systems

13.1 Introduction

Linear-response time-dependent density functional theory (TDDFT) [250, 251, 306] is very widely used to model electronic excited states of chemical species. TDDFT is an appealing approach as it is computationally inexpensive ($O(N^{3-4})$ scaling vs number of basis functions N), nearly black-box and able to simultaneously compute a large number of excited states. However, the lack of explicit orbital relaxation renders it unsuitable for describing excitations that involve substantial reorganization of electron density, such as charge transfer [251, 387] or Rydberg excited states [325, 406]. Excitation of core electrons in particular involves a substantial relaxation of the core-hole (and an accompanying reorganization of valence electron density), which leads to substantial errors in excitation energies predicted by TDDFT by standard functionals. It is consequently not unusual to blue-shift TDDFT core-level spectra by ~ 10 eV for alignment with experiment [328, 460–462, 522, 523](though the qualitative nature of transitions is typically reasonably predicted). Some specialized short-range corrected functionals specifically trained to predict core-level spectra [377] tend to fare better [524–530], but the very strong sensitivity of TDDFT excitation energies on delocalization error [64, 251] is troubling (as even small perturbations could have disproportionate impact on relative peak positions).

In contrast, linear-response based wavefunction theories like equation of motion coupled cluster singles and doubles (EOM-CCSD) [381, 382, 531–536] tend to systematically overestimate core-excitation energies [378, 465, 532, 537] due to lack of explicit orbital relaxation, often necessitating empirical redshifting by 1-2 eV for alignment with experiment [378, 532, 533, 537]. Encouragingly however, use of different core-valence separation [538] (CVS) schemes has been observed to reduce the magnitude of the shift required [536, 539, 540]. A flavor of second order extended algebraic diagrammatic construction (ADC(2)-x [466], specifically CVS-ADC(2)-x [460]) is also employed to calculate core-level spectra [460, 541–543].

The accuracy of CVS-ADC(2)-x however owes a great deal to fortuitious cancellation between various sources of error [460, 544], and performance actually worsens when third order ADC is employed [544]. At any rate, the higher computational cost of these wavefunction theories ($O(N^6)$ for both EOM-CCSD and ADC(2)-x) and slower basis set convergence renders them impractical for large molecular systems or extended materials, relative to computationally inexpensive DFT approaches. Nonetheless, development of lower-scaling approximations to these wavefunction based methods is expected to broaden their applicability considerably [532, 545].

In contrast to these linear-response based protocols, state specific orbital optimized (OO) methods have been much more successful at accurate prediction of core-level spectra even within the DFT paradigm [416, 474, 475, 546, 547]. The main difficulty with these methods is the potential for 'variational collapse' of the target excited state down to the ground state or another excited state, as it is challenging to optimize excited state orbitals (by virtue of excited states typically being saddle points of energy). The maximum overlap method (MOM) [341, 344] was developed to address this problem for repeated Fock matrix diagonalization based methods like DIIS [342], though convergence failures and variational collapse (via slow drifting of orbitals) are not always prevented [343, 345]. More recently, some of us have have proposed a square gradient minimization (SGM) [345] based direct minimization approach that appears to be robust against both modes of MOM failure. SGM has been employed in conjunction with the spin-pure restricted open-shell Kohn-Sham (ROKS) method [362, 363] to predict highly accurate ($<$ 0.5 eV error) core-level spectra of closed-shell molecules [546] at local DFT cost (using the modern SCAN [62] functional). It is also worth noting that there exist linear-response methods that incorporate partial OO character through relaxed core-ionized states, like Static Exchange (STEX) [476] or Non-orthogonal Configuration Interaction Singles (NOCIS) [376, 477, 496], though such treatments are wavefunction based and $\sim$ 1 eV error remains common due to lack of dynamic correlation.

Stable open-shell molecules are fairly uncommon in nature and there is consequently a scarcity of static experimental spectra for such species. However open-shell systems are omnipresent in chemical dynamics experiments (either as fragments or excited states of closed-shell molecules) where transient X-ray absorption spectroscopy (XAS) is often employed [373–375, 548]. It is consequently useful to have cheap and reliable theoretical techniques capable of modeling core-level spectra of such species. The highly accurate ROKS method is however not applicable to most open-shell systems, as it is explicitly designed for singlet states with one broken electron pair. In fact, open-shell systems pose additional challenges for many of the methods described above, as a spin-pure treatment of excited states necessitates inclusion of some double excitations [322, 477, 496] even for states that conventionally appear to be single excitations breaking one electron pair. This is not too difficult for wavefunction approaches, as shown by the extended CIS (XCIS [322]) and open-shell NOCIS [477, 496] methods. However, it is not at all straightforward to achieve this within TDDFT, which has no route for describing double excitations within the widely used adiabatic approximation [251, 326, 327]. It is tempting to believe that missing such configurations would not be particularly significant if the unpaired electrons interact only weakly, but the failure of TDDFT in describing excited state single bond dissociations despite the unrestricted reference state being reasonable [228] indicates some cause for caution.

In this work, we apply OO excited state DFT in conjunction with SGM to study single core-excitations of open-shell systems. This entails investigation of excitations to both singly occupied levels (which can be well described by single determinants, in principle) and completely unoccupied levels (which result in intrinsically multiconfigurational states). We present a scheme for recoupling multiple configurations to obtain an approximate doublet state for the latter class of excitations and discuss where this protocol might be necessary by considering the C K-edge spectra of the allyl radical, O K-edge of CO^+ and the N K-edge of NO_2. We also discuss general principles for reliably using these techniques to predict core-excitation spectra. Overall, we demonstrate that highly accurate DFT results can be obtained via orbital optimization with the modern local SCAN functional at low computational cost, similar to behavior observed for closed-shell systems. Low error can also be achieved via cam-B3LYP, TPSS and ωB97X-D3 functionals (albeit at a somewhat higher asymptotic cost for the hybrid functionals).

13.2 Theory

13.2.1 Single Determinant States

Excitations from the core to singly occupied molecular orbitals (SOMOs) of open-shell systems result in states representable via a single Slater determinant, as there is no change in the number of unpaired electrons. The simplest approach for modeling such states is Δ Self-Consistent Field (ΔSCF) [264, 341, 358, 416], where the non-aufbau solution to the Hartree-Fock [14] or Kohn-Sham [48] DFT equations is converged via an excited state solver like SGM or MOM. The resulting excited KS determinant would not necessarily be exactly orthogonal to the ground state determinant but this is generally of little concern since KS determinants are fictitious entities useful for finding densities and thus there exists no requirement that ground and excited state determinants be orthogonal. Nonetheless, a significant (> 0.1, for example) squared overlap between the ground and excited state configurations would be concerning but we have not observed such occurrences in our investigations and do not believe them to be likely without at least partial variational collapse of the core-hole.

The principal dilemma for such states is choosing between spin-restricted or unrestricted orbitals for ΔSCF. Unrestricted orbitals are typically more suitable for DFT studies on open-shell systems, though some functionals are known to yield atypically unphysical behavior in certain limits away from equilibrium [218]. On the other hand, restricted open-shell (RO) orbitals artificially enforce a spin-symmetry that does not exist in radicals. As will be shown later (in Table 13.2), use of unrestricted orbitals appears to systematically lower the core-excitation energies (via extra stabilization of the core-excited state relative to the ground state). The best functionals for predicting spectra of closed-shell species yield lower errors for radicals when unrestricted orbitals are employed, and we thus recommend the use of unrestricted orbitals over RO orbitals for radicals. RO orbitals however should be employed for closed-shell systems (via ROKS or related methods) [546], on account of the existence of spin-symmetry in such species.

13.2.2 Multiconfigurational States

Multiconfigurational DFT is a difficult challenge even outside the unique challenges of TDDFT for double excitations, as the Kohn-Sham (KS) exchange-correlation energy is defined for a single determinant reference. KS-DFT target states therefore should be single determinants, and directly recoupling them via configuration interaction (CI) would result in double counting of some electron-electron interactions through both the functional and the CI off-diagonal terms. This is quite undesirable, making modeling such states fairly challenging.

One very reasonable solution is to note that single determinants with both α and β unpaired electrons are mixtures of different spin-states, and the highest spin-state within that ensemble can be well approximated by a single determinant by merely making all unpaired spins point in the same direction. Approximate spin-projection (AP) [357] can consequently be applied to remove this high spin contribution from a spin impure mixed determinant. This approach should be sufficient when there are only two significantly contributing eigenstates to the mixed configuration, as is the case for single excitations out of closed-shell molecules (where only the singlet and triplet states contribute). ROKS in fact utilizes this very feature to ensure spin-purity. ROKS employs a mixed configuration that has one unpaired α spin and one unpaired β spin (which has energy E_M) and a triplet configuration that has both unpaired spins as α (which has energy E_T). The use of RO orbitals forces the mixed configuration to be exactly halfway between singlet and triplet, indicating $E_M = \dfrac{E_S + E_T}{2}$ where E_S is the true singlet energy. ROKS consequently optimizes the purified singlet energy $E_S = 2E_M - E_T$.

Things are however substantially more challenging for doublet states. A mixed configuration with two unpaired α electrons and one unpaired β electron is a mixture of three states—two doublets and a quartet. The quartet contribution can be easily removed using an AP protocol similar to ROKS, but disentangling the two doublet energies is nontrivial.

Looking at the pure wavefunction based CI approach however offers some hints as to how to proceed. If we consider restricted open-shell configurations with three unpaired electrons occupying three spin-restricted orbitals (labeled $1, 2$ and 3, respectively), eight possible configurations exist. Spin-inversion symmetry in the absence of magnetic fields however indicate that only four provide unique information:

1. $|Q\rangle = |\uparrow\uparrow\uparrow\rangle$: All three spins are α. This is the pure quartet with energy E_Q.
2. $|M_1\rangle = |\downarrow\uparrow\uparrow\rangle$: Only the spin at orbital 1 is β. This is a mixed configuration with energy $E_{M_1} = E_Q + K_{12} + K_{13}$, where K_{pq} is the exchange interaction $\langle pq|qp\rangle$ between an electron in orbital p and another in orbital q. The inversion of the spin in orbital

1 relative to the quartet leads to a loss of exchange stabilization between this orbital and the other two, leading to the energy going up by $K_{12} + K_{13}$.

3. $|M_2\rangle = |\uparrow\downarrow\uparrow\rangle$ Only the spin at orbital 2 is β. Consequently $E_{M_2} = E_Q + K_{12} + K_{23}$.
4. $|M_3\rangle = |\uparrow\uparrow\downarrow\rangle$ Only the spin at orbital 3 is β. Consequently $E_{M_3} = E_Q + K_{13} + K_{23}$.

Having the single determinant energies $E_Q, E_{M_1}, E_{M_2}, E_{M_3}$ is sufficient to uniquely solve for the exchange interactions K_{pq}, with $K_{12} = \dfrac{E_{M_1} + E_{M_2} - E_Q - E_{M_3}}{2}$ etc. This is quite useful, as the off-diagonal CI coupling elements are $\langle M_i| H |M_j\rangle = -K_{ij}$ from Slater-Condon rules for double excitations [14]. This indicates that the knowledge of the single determinant energies is sufficient for solving the CI problem. With this, we find the eigenvalues of H within the subspace spanned by $|M_{1,2,3}\rangle$ to be:

$$E_1 = E_Q \tag{13.1}$$

$$E_2 = \frac{1}{2}\left(E_{M_1} + E_{M_2} + E_{M_3} - E_Q - \sqrt{2\left[(E_{M_1} - E_{M_2})^2 + (E_{M_2} - E_{M_3})^2 + (E_{M_3} - E_{M_1})^2\right]}\right) \tag{13.2}$$

$$E_3 = \frac{1}{2}\left(E_{M_1} + E_{M_2} + E_{M_3} - E_Q + \sqrt{2\left[(E_{M_1} - E_{M_2})^2 + (E_{M_2} - E_{M_3})^2 + (E_{M_3} - E_{M_1})^2\right]}\right) \tag{13.3}$$

The first eigenvalue corresponds to the quartet within the $M_S = \frac{1}{2}$ subspace (which is a linear combination of all three configurations with equal weights). The other two correspond to the energies of the two possible doublet states.

We propose that the same approach be employed for recoupling DFT configurations, with the KS energies of configurations $|M_{1,2,3}\rangle$ being employed instead of the HF ones used in the wavefunction theory approach. The risk of double counting should be greatly reduced as the effective off-diagonal elements are found directly from the KS energies versus Slater-Condon rules. Indeed, the off-diagonal elements should no longer be viewed as exchange interactions but rather effective spin-spin coupling elements. The entire approach is basically equivalent to solving for the eigenstates of the effective Ising like Hamiltonian $H' = -2J_{12}\vec{S}_1 \cdot \vec{S}_2 - 2J_{13}\vec{S}_1 \cdot \vec{S}_3 - 2J_{23}\vec{S}_2 \cdot \vec{S}_3$ for three interacting spins, where the couplings J_{ij} are obtained from DFT (and are equivalent to the exchange interactions K_{ij} if HF is used as the functional). Such approaches have been used within broken-symmetry DFT to calculate spin coupling constants of transition metal species to reasonable accuracy for instance [549–556], and it is hoped that similar behavior will transfer over. Furthermore, following the equivalent logic for the case of two unpaired spins yields ROKS, which is known to be quite accurate for such states [265, 345, 363]. These known instances of successful behavior encourages us to believe that this protocol is worthwhile to explore. We also note that Eqns 13.2-13.3 were reported in Ref 358 without an explicit description of the derivation, but these have not been actually applied to core-level spectroscopy (or any excited state problem) to the best of our knowledge.

Having obtained $E_{2,3}$ as spin-purified energies, we next seek to determine how to obtain the optimal orbitals. It is tempting to directly optimize $E_{2,3}$ in a manner analogous to ROKS but we have elected not to do so at present. This optimization is nontrivial due to the nonlinear nature of the energy expression (vs the simpler form for ROKS). In addition, the derived equation is only precisely true for restricted open-shell orbitals, while Sec 13.2.1 seems to suggest unrestricted orbitals are optimal. We therefore look to AP-ΔSCF [264, 358] for singlet excited states for inspiration, where the mixed determinant and triplet determinants are individually optimized (resulting in two sets of orbitals) and the singlet energy is simply computed as $2E_M - E_T$ from the individually optimized energies, instead of optimizing a single set of orbitals as in ROKS. The resulting energies however are often not dramatically different from ROKS [363] and so we choose to follow a similar protocol here to determine if there is sufficient utility in this route for recoupling mixed determinants to justify optimizing a single set of unrestricted orbitals for computing the doublet energies. We consequently optimize $|Q\rangle$ and $|M_{1,2,3}\rangle$ individually and compute $E_{2,3}$ from those optimized energies.

One rather inconvenient detail is that individually optimized $|M_{1,2,3}\rangle$ configurations would thus not be strictly orthogonal to each other due to slight differences in the orbitals. However we do not consider any non-orthogonality derived terms arising from mixed configurations, as the KS determinants are fictitious constructs. On a more practical note, we ensure low overlap via providing restricted open-shell quartet orbitals as the initial guess for SGM optimization of the mixed determinants. The initial guesses are thus orthogonal, and orbital relaxation to the closest stationary point (which SGM is supposed to achieve) in unrestricted space should not lead to significant non-orthogonality for cases where this model of three unpaired electrons is a good approximation. Further details about initial guesses are enumerated in Sec 13.4.

We also note that the work here focused only on the case of three unpaired electrons. Ref 366 presents the equations for the four unpaired electron case, as well as a discussion on the general approach.

13.2.3 Transition Dipole Moments

The magnitude of the transition dipole moment between the ground and excited states is essential for computing oscillator strengths (and thus relative intensities in computed spectra). The fictitious nature of the KS determinant (which represents a wavefunction of noninteracting electrons subjected to a fictitious potential) is a significant obstacle here, as it implies there is no rigourous route for computing transition dipole moments. However, treating the KS determinants as real wavefunctions might be a reasonable approximation for computing this quantity, in the hope that the KS determinants (or superpositions thereof) would have a reasonably large overlap with the true wavefunctions to make this exercise worthwhile. Indeed, spectra computed via this route show fairly good agreement with experiment (as can be seen from previous work [546] by some of us, for instance). Such a protocol can (and should) account for nonorthogonality between ground and mixed determinants as it is fairly simple to compute NOCI dipole matrix elements [394].

There are some additional factors to consider for the recoupled multiconfigurational states. The wavefunction inspired approach indicates that transition dipole moments should be computed via a linear combination of the transition dipole moments of individual determinants, as weighted by their coefficients in the eigenvectors corresponding to $E_{2,3}$. The effect of non-orthogonality between mixed determinants $|M_{1,2,3}\rangle$ on eigenvector coefficients is neglected here both because such terms are relatively small (because the mixed determinants have fairly low overlap with each other) and because it is not straightforward to calculate these effects. The decision to not consider this form of nonorthogonality does not appear to have any significant deleterious impact, as shown by the spectra presented later.

The other important factor to consider is that the analysis in Sec 13.2.2 found off-diagonal coupling elements directly from the energies $E_{M_{1,2,3}}$ and thus did not account for phases of $|M_{1,2,3}\rangle$. These phases however are critical for estimating transition dipole moments, and thus must be obtained somehow. A protocol for estimating these phases via the formally "quartet" state is supplied in Sec 13.7.

13.3 Results and Discussion

13.3.1 Excitations to the SOMO

The relative scarcity of experimental XAS data for radicals leaves us with a fairly small dataset of 17 excitations for assessing the performance of single determinant ΔSCF. The precise statistical values here are thus less reliable than those obtained in Ref 546 from 40 excitations out of closed-shell molecules, but general qualitative trends can be drawn even from this restricted amount of data. The experimental excitation energies for all the C K-edge excitations (save allyl and CO^+) were measured by some of us, via radicals obtained from the photodissociation of the corresponding iodide [548, 557]. These values should have an uncertainty of ±0.1 eV, although vibrational excitations induced by photodissociation could shift the values somewhat. However, the resulting excitation energy for CH_3 agrees well with vibrationally resolved spectra obtained from radicals generated from flash pyrolysis [558]. Furthermore, (as can be seen from Table 13.1), the experimental shifts between the C K-edge of the allyl radical (obtained by authors of Ref 559 on cold radicals generated via flash pyrolysis) and other C K-edges are very well reproduced by theoretical methods, suggesting that any vibrational excitation induced effect was small overall. A full Frank-Condon analysis could prove useful in quantifying any such effect, but was not pursued at present.

We only consider a relatively small number of density functionals as a combination of large experimental uncertainty (typically 0.1 eV) and limited number of data points would make precise rankings of many functionals meaningless. We think it is more useful to investigate the performance of some representative functionals and see if they are sufficiently accurate to justify wider use. We therefore consider the following functionals from various rungs of Jacob's ladder [64]:

1. Rung 1 (local spin-density approximation/LSDA [54, 138, 149]): Not considered due to very large errors found in Ref 546.
2. Rung 2 (generalized gradient approximation/GGA): BLYP [58, 151], PBE [57].
3. Rung 3 (meta-GGA): TPSS [61], SCAN [62].
4. Rung 4 (hybrids): B3LYP [89], PBE0 [91] (global hybrids); cam-B3LYP [94], ωB97X-D3 [167], ωB97X-V [108] (range separated hybrids).
5. Rung 5 (double hybrids): Not considered due to significant computational expense.

The Hartree-Fock (HF) wavefunction method is also considered, in order to determine the impact of neglecting correlation entirely. The choice of functionals here was guided by both a desire to compare against closed-shell results reported earlier [546] and a desire to examine the behavior of classic, minimally parameterized functionals like B3LYP.

Radical	Expt.	BLYP	PBE	TPSS	SCAN	B3LYP	PBE0	cam-B3LYP	ωB97X-D3	ωB97X-V	HF
CH_3	281.4 [558]	281.6	280.8	281.8	281.8	281.7	281.2	281.6	281.8	281.9	282.8
CH_3CH_2	281.7 [557]	282.0	281.3	282.1	282.2	282.1	281.6	282.0	282.2	282.3	283.0
$(CH_3)_2CH$	282.2 [557]	282.3	281.6	282.4	282.5	282.4	282.0	282.3	282.5	282.6	283.3
$(CH_3)_3C$	282.6 [557]	282.6	281.9	282.6	282.8	282.7	282.3	282.6	282.8	282.9	283.5
Allyl	282.0 [559]	282.3	281.5	282.4	282.5	282.4	281.9	282.3	282.5	282.6	283.5
CO^+	282.0 [560]	282.2	281.3	282.2	282.3	282.3	281.8	282.2	282.4	282.5	283.3
CH_2Br	282.6 [557]	282.8	282.0	282.9	282.9	282.8	282.4	282.7	282.9	283.1	283.8
CH_2Cl	282.8 [548]	283.0	282.2	283.1	283.2	283.0	282.6	282.9	283.1	283.3	284.0
NH_2	394.3 [561]	394.5	393.6	394.6	394.7	394.5	394.1	394.5	394.7	394.8	395.7
N_2^+	394.3 [562]	394.4	393.5	394.3	394.3	394.2	393.6	394.0	394.2	394.3	394.4
NH_3^+	395.2 [563]	395.0	394.2	395.2	395.3	395.1	394.6	395.0	395.2	395.4	396.4
NO_2	401.0 [564]	401.0	400.2	401.0	401.2	401.0	400.6	401.1	401.3	401.5	402.1
OH	525.8 [565]	525.8	524.9	526.0	526.0	525.8	525.3	525.8	525.9	526.1	527.0
HO_2	528.6 [566]	528.5	527.6	528.6	528.5	528.3	527.8	528.3	528.4	528.6	528.9
NO_2	530.3 [564]	530.5	529.7	530.5	530.5	530.2	529.7	530.1	530.2	530.4	529.9
O_2	530.8 [567]	530.8	530.0	530.8	530.8	530.6	530.1	530.6	530.8	531.0	530.6
CO^+	528.5 [560]	528.3	527.5	528.4	528.5	528.0	527.6	528.0	528.1	528.2	528.3
	RMSE	0.2	0.7	0.2	0.3	0.3	0.5	0.2	0.3	0.4	1.1
	ME	0.1	-0.7	0.2	0.2	0.0	-0.4	0.0	0.2	0.3	0.9
	MAX	0.3	1.0	0.4	0.5	0.5	0.9	0.5	0.5	0.6	1.5

Table 13.1: ΔSCF/aug-cc-pCVTZ [70, 198, 437] core to SOMO excitation energies (in eV) for open-shell species, as predicted by various functionals. Unrestricted orbitals were used for both the ground and excited states. Root mean squared error (RMSE), mean error (ME) and maximum absolute error (MAX) are also reported.

Table 13.1 presents the excitation energies calculated using the chosen approaches (using spin-unrestricted orbitals), along with statistical measures of error like the root mean squared error (RMSE). None of the density functional methods deviate from experiment by more than 1 eV, which is in sharp contrast to the typical behavior of TDDFT with the same functionals [546]. Even HF has only < 2 eV error despite complete absence of correlation. We specifically observe that the BLYP, TPSS, SCAN, B3LYP, cam-B3LYP and ωB97X-D3 functionals yield 0.3 eV or lower RMSE, and do not deviate by more than 0.5 eV from the experimental reference values. The good performance by local functionals like BLYP, TPSS and SCAN is quite impressive, as these functionals are much more computationally efficient than hybrids. Of the trio, the performance of only SCAN has been characterized for closed-shell systems [546], where it was also found to be similarly accurate. We consequently focus on the performance of SCAN in later sections of this work, as good performance in both the closed and open-shell limits is critical for prediction of transient X-ray absorption spectroscopy. However, we believe that good performance can be obtained from many functionals considered in this work (as partially demonstrated in the Supporting Information). Interestingly, PBE and PBE0 perform surprisingly poorly, especially relative to BLYP and B3LYP, respectively.

Table 13.1 furthermore shows that the small errors for many functionals are mostly systematic, which appears to suggest that the change in excitation energy between two species (say between methyl and tert-butyl, for instance) would be reproduced fairly accurately by most functionals. This is also in principle true for TDDFT, although the massive (~ 10 eV) errors in the individual excitation energies mean that even a relatively small variation in absolute error could have significant impact on relative peak positions (made more likely by very high sensitivity of TDDFT results to delocalization error [64, 251]). Most functionals (including SCAN) appear to systematically overestimate energies, while PBE and PBE0 systematically underestimate (which might be the reason for their poor overall performance). Inclusion of relativistic effects [495] (which systematically increase excitation energies by binding core electrons more tightly) would therefore degrade performance of many functionals, while improving the performance of PBE and PBE0. The atom specific relativistic corrections for C,N and O are however quite small [495] (0.1-0.3 eV) and therefore are often neglected in studies (such as by the SRC functionals trained for TDDFT spectra prediction [377], which has these effects implicitly baked into what is fundamentally a non-relativistic theory). The impact of incorporating these corrections on the errors of various models is provided in the supporting information, which shows that the RMSE of functionals (other than PBE and PBE0) goes up by 0.1 eV at most, suggesting that this is not a major issue in practice. We also note that HF systematically overestimates excitation energies by ~ 1 eV due to missing correlation, which indicates that simple models for dynamical correlation (such as perturbative approaches [14, 19]) might be adequate for substantially lowering error, albeit at higher computational cost than DFT. HF however has a strong propensity to spuriously spin-contaminate Slater determinants, and the performance of perturbative corrections to HF references could consequently be greatly degraded [19, 568].

We also consider whether there is any benefit to using restricted open-shell orbitals over

Radical	Experiment	RO-SCAN	USCAN	RO-PBE0	UPBE0
CH_3	281.4	281.9	281.8	281.3	281.2
CH_3CH_2	281.7	282.3	282.2	281.7	281.6
$(CH_3)_2CH$	282.2	282.6	282.5	282.1	282.0
$(CH_3)_3C$	282.6	282.9	282.8	282.4	282.3
Allyl	282.0	282.5	282.5	281.9	281.9
CO^+	282.0	282.5	282.3	281.9	281.8
CH_2Br	282.6	283.0	282.9	282.5	282.4
CH_2Cl	282.8	283.3	283.2	282.7	282.6
NH_2	394.3	394.8	394.7	394.2	394.1
N_2^+	394.3	394.5	394.3	393.8	393.6
NH_3^+	395.2	395.4	395.3	394.7	394.6
NO_2	401.0	401.4	401.2	400.7	400.6
OH	525.8	526.2	526.0	525.4	525.3
HO_2	528.5	528.6	528.5	527.7	527.6
NO_2	528.6	528.7	528.5	527.9	527.8
O_2	530.3	530.7	530.5	529.8	529.7
CO^+	530.8	531.0	530.8	530.2	530.1
	RMSE	0.4	0.3	0.4	0.5
	ME	0.4	0.2	-0.3	-0.4
	MAX	0.6	0.5	0.8	0.9

Table 13.2: Comparison of ΔSCF/aug-cc-pCVTZ core to SOMO excitation energies (in eV) for restricted open-shell (RO) and unrestricted (U) orbitals with SCAN and PBE0. Results for other functionals are provided in the supporting information.

unrestricted orbitals. Table 13.2 indicates that use of unrestricted orbitals systematically lowers excitation energies by ~ 0.1 eV relative to restricted open-shell results. This consequently indicates that use of RO orbitals instead of U would degrade the performance of most of the studied functionals (as they systematically overestimate with U orbitals) and improve the behavior of PBE and PBE0. Indeed, Table 13.2 shows that both RO-PBE0 and RO-SCAN have the same RMSE of 0.4 eV. This potentially argues that RO-PBE0 is perhaps preferable to USCAN, as the small relativistic corrections furthers improve the RO-PBE0 RMSE to 0.2 eV (while degrading USCAN's RMSE to 0.4 eV, as shown in the Supporting Information). However, we believe that SCAN with unrestricted orbitals is still the preferred route, even aside from greater asymptotic computational efficiency. Open-shell systems tend to often arise in transient absorption experiments starting from closed-shell species, and so it is important to use an approach that is effective at predicting the spectra for both types of systems. PBE0 is perceptibly inferior to SCAN when it comes to closed-shell systems [546] (irrespective of inclusion of relativistic effects), and the two are fairly close in predictive ability for open-shell systems, making SCAN with unrestricted orbitals the preferred choice. We also note that a comparison between aug-cc-pCVTZ and aug-cc-pCVQZ results shows that a small part ($\sim$ 0.1 eV) of the systematic overestimation predicted by SCAN for Table 13.1 values stems from basis set incompleteness (as shown by a comparison in the Supporting Information), similar to behavior of closed-shell species [546].

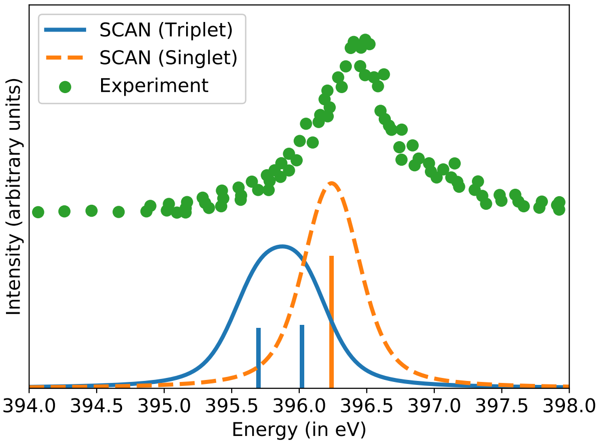


Figure 13.1: Comparison of experimental N K-edge spectrum of NH_2^+ (obtained from Ref 563) with those computed with SCAN/aug-cc-pCVTZ. A Voigt profile with a Gaussian standard deviation of 0.2 eV and Lorentzian $\gamma = 0.121$ eV was utilized for broadening the computed spectra. Bars are supplied to denote the location of the predicted excitation energies. The singlet and triplet spectra have been normalized by the same factor for a fair comparison.

13.3.1.1 The Case of NH_2^+

The spectrum of NH_2^+ has been experimentally characterized [563], but was not considered in Table 13.1 as the two possible excitations to singly occupied levels are unresolved experimentally (assuming the radical cation is in the ${}^3B_1$ ground state). We used ΔSCF to compute the two transitions separately, and report them in Table 13.3. These transitions have nearly the same oscillator strength and thus their average should roughly correspond to the experimental peak. The ROKS results for the lowest lying singlet ${}^1A_1$ excited state is also reported, in case it contributes to the experimental spectrum as well. Fig 13.1 presents the representative case of the SCAN functional, with other methods yielding similar figures.

The computed average triplet excitation energies in Table 13.1 agree fairly well with experiment, especially for good performers like SCAN, B3LYP or ωB97X-D3. However, the values are somewhat red-shifted, in stark contrast to the general behavior seen in Table 13.1. One possible explanation for this would be a blue-shifting of the experimental spectrum due to presence of singlet NH_2^+, since this state absorbs fairly strongly (roughly twice the oscillator strength than the individual triplet transitions) at slightly higher energies than the triplet, pushing the overall center of the band to higher energies (as hinted at by Fig 13.1). However, the computed triplet excitation average and the experimental maximum

Method	Triplet			Singlet (ROKS)
	Low	High	Average	
Experiment			396.4 [563]	
BLYP	395.7	396.0	395.8	396.1
PBE	394.9	395.2	395.0	395.3
TPSS	396.0	396.3	396.2	396.2
SCAN	395.7	396.0	395.9	396.2
B3LYP	395.7	396.0	395.9	396.1
PBE0	395.3	395.6	395.4	395.7
cam-B3LYP	395.7	396.0	395.9	396.1
ωB97X-D3	396.1	396.4	396.2	396.2
ωB97X-V	396.3	396.6	396.4	396.3

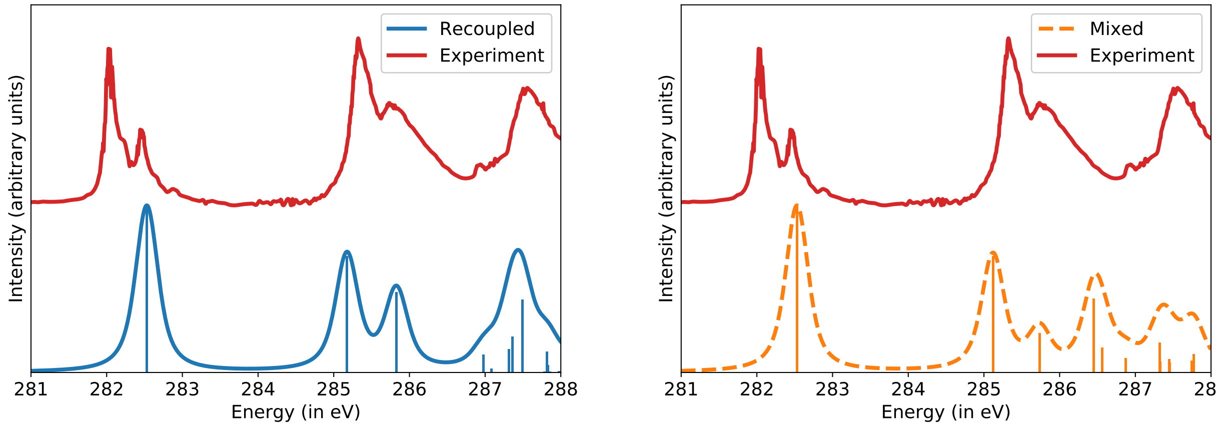
Table 13.3: Comparison of ΔSCF/aug-cc-pCVTZ core to SOMO excitation energies (in eV) for $^3B_1$ NH_2^+ with various functionals. The lowest core excitation energy for the $^1A_1$ singlet is also reported

are not too far from each other (0.5 eV for SCAN), so it is not entirely impossible for DFT error to be the sole reason behind the discrepancies. ωB97X-V for instance gives quite good agreement with experiment, without needing to invoke the singlet state.

13.3.2 Spectrum of the Allyl Radical

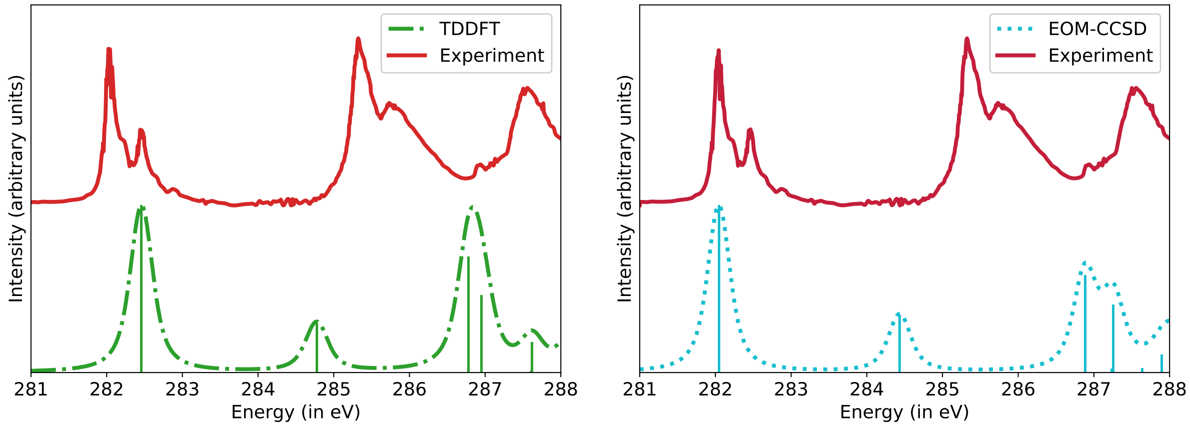
Having explored the utility of ΔSCF in predicting excitation energies to the SOMO, we next seek to investigate the utility of the theory described in Secs 13.2.2 and 13.2.3 at predicting the full core-excitation spectrum. The recoupling approach described therein is expected to be most effective for excitations to unoccupied valence orbitals, as then all three unpaired spins (in the core, SOMO and valence excited levels) will be interacting strongly. The lack of experimental spectra to compare against is again a problem, and restricts us to only a few data points. Fortunately, the allyl radical has an experimentally characterized spectrum [559] that is dominated by excitations to the unoccupied π^* LUMO orbital, making it an excellent example for determining the utility of our recoupling approach, relative to simply using mixed configurations alone.

Fig 13.2 compares the performance of the orbital optimized methods in reproducing the C K-edge spectrum of the allyl radical. The performance of fc-CVS-EOM-CCSD and TDDFT with the specialized short-range corrected SRC2-R1 [377] functional is also considered. All three DFT methods are reasonable at predicting the lowest energy allowed excitation (from the terminal C atoms to the SOMO, the corresponding transition from the central C atom being symmetry forbidden), though all systematically overestimate by approximately 0.5 eV, resulting in the computed peak aligning with the vibrational fine structure of the experimental band. This is potentially indicative of some multireference character of this excited state,



(a) Recoupled configurations/SCAN

(b) Mixed Configurations/SCAN



(c) TDDFT/SRC2-R1

(d) fc-CVS-EOM-CCSD [536]

Figure 13.2: Comparison of experimental C K-edge spectrum of the allyl radical (obtained from Ref 559) with those computed with DFT/aug-cc-pCVTZ and fc-CVS-EOM-CCSD/aug-cc-pCVTZ. The SRC2-R1 functional was employed for the TDDFT spectrum, while SCAN was utilized for both the recoupled and mixed configuration approaches. A Voigt profile with a Gaussian standard deviation of 0.1 eV and Lorentzian $\gamma = 0.121$ eV was utilized for broadening the computed spectra. Bars are supplied to denote the location of the predicted excitation energies.

though it is difficult to draw firm conclusions from density functional data alone (especially since it is possible to get better agreement via a functional that systematically underestimates 1s→SOMO excitation energies, like PBE0). It is however worth noting that fc-CVS-EOM-CCSD is spot on for this excitation, without any need for empirical translation of spectrum (as can be seen from Fig 13.2d).

Fig 13.2c lays bare the failure of TDDFT at predicting excitations to the LUMO, as the peak positions are completely off. This is not a pecularity of the SRC2-R1 functional but rather a failure of the TDDFT family of methods, as translated TDDFT spectra from other functionals yield a similarly poor picture (as shown in the Supporting Information). Similarly, Fig 13.2d shows that fc-CVS-EOM-CCSD is unable to yield a qualitatively better spectrum than TDDFT, further highlighting the inadequacies of linear-response methods for this system. It is indeed somewhat puzzling that the inclusion of double excitations in fc-CVS-EOM-CCSD did not lead to any significant improvement over TDDFT (which is restricted to single excitations alone). The qualitative failure of both linear-response methods is likely a consequence of both spin-contamination and lack of orbital relaxation. Explicit inclusion of triple excitations could potentially ameliorate both issues but the significant computational expense of full EOM-CCSDT would dramatically constrain practical use.

The SCAN based orbital optimized approaches fare better, with both spin-contaminated mixed determinants and the recoupling approach yielding roughly qualitatively correct behavior. However, Fig 13.2b shows that the mixed determinant approach fails to accurately predict the energy of the higher energy central C to LUMO transition, underestimating it by an eV. This substantially damages the quality of the predicted spectrum, by making this peak appear in an area where none are present experimentally.

Bright Transitions	Experiment	MCSCF	Recoupled	Mixed	TDDFT	EOM-CCSD
			SCAN	SCAN	SRC2-R1	
C_T →SOMO	282.0	281.9	282.5	282.5	282.5	282.0
C_C →LUMO	285.3	285.7	285.2	285.1	284.8	284.4
C_T →LUMO	285.7	285.9	285.8	285.7	287.0	287.3
C_C →LUMO	287.5	288.3	287.5	286.5	286.8	286.9

Table 13.4: Comparison of experimentally observed excitation energies (in eV) in the allyl core absorption spectrum with theoretical methods. The experimental values and MCSCF numbers were obtained from Ref 559. C_T is terminal carbon, while C_C is central carbon.
.

The recoupling approach shifts this peak to the appropriate location and predicts a spectrum in excellent agreement with experiment (as can be seen from Fig 13.2a). Indeed, Table 13.4 shows that the peaks predicted by recoupled SCAN agree better with experiment than MCSCF calculations reported in Ref 559 (though not too much should be inferred from this single data point). This good performance is not unique to SCAN alone, as several other functionals yield similar spectra in both the recoupled and mixed regimes (as shown in the Supporting Information). Specifically, we find that recoupled cam-B3LYP, PBE0 and TPSS give good predictions for the 1s→ LUMO portion of the spectrum, while BLYP and PBE yield rather poor performance even after recoupling. SCAN and cam-B3LYP appear to give the best performance, while some of the higher energy peaks with PBE0 and TPSS are somewhat redshifted with respect to the experimental spectrum. This supports our decision of selecting SCAN as the principal functional for the manuscript, despite BLYP and TPSS having the same computational scaling and slightly lower RMSE for excitations to SOMO (as shown in Table 13.1). The poor qualitative performance by BLYP and PBE also serves as a potential warning against attempting to use GGAs for prediction of core spectra, despite BLYP's excellent behavior for excitations to SOMO. Ultimately, the recoupling scheme cannot correct for deficient physics in the mixed configuration energies and a poor choice of functional could lead to poor results. Nonetheless, it is encouraging to see that all 'advanced' functionals (Rungs 3 and 4) tested yield a reasonable spectrum after recoupling.

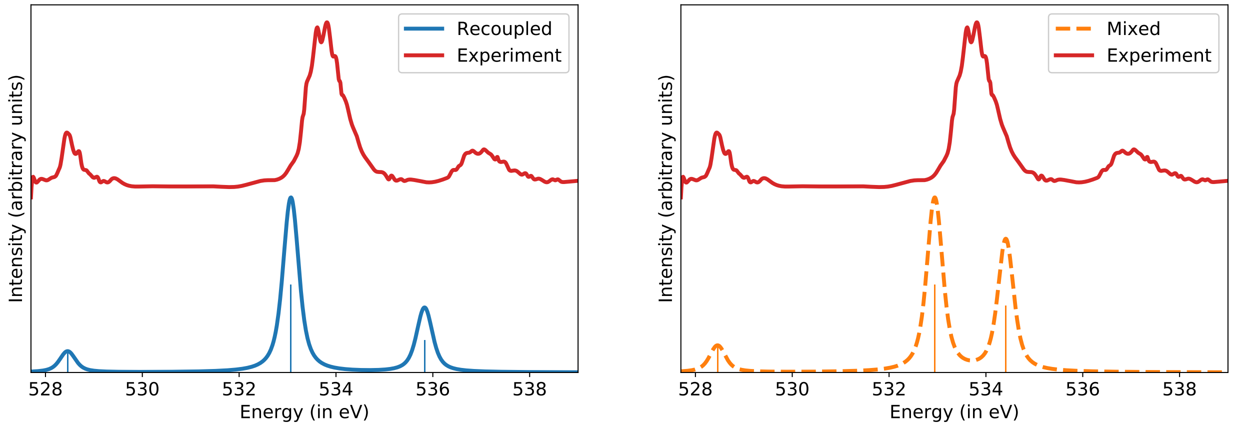
Overall, this example seems to suggest that orbital optimized approaches have an edge over TDDFT/EOM-CCSD when it comes to predicting core-excitation spectra of radicals. Furthermore, recoupling spin-contaminated mixed configurations to yield approximate doublets appears to not degrade performance and leads to some improvements. The overall accuracy of recoupled SCAN at predicting the spectrum of allyl certainly appears to hint at the efficacy of using this approach for XAS studies of large carbon based polyradical systems, such as ones that might arise in soot formation during combustion [569].

13.3.3 O K-edge Spectrum of CO^+

We next consider the rather challenging case of the CO^+ radical cation, whose experimental spectrum has been characterized very recently [560]. We focus on the O K-edge as the two doublet states corresponding to the 1s→LUMO excitation are experimentally well resolved, unlike the C K-edge (where vibrational fine structure of the lower energy excitation overlaps with the higher energy one).

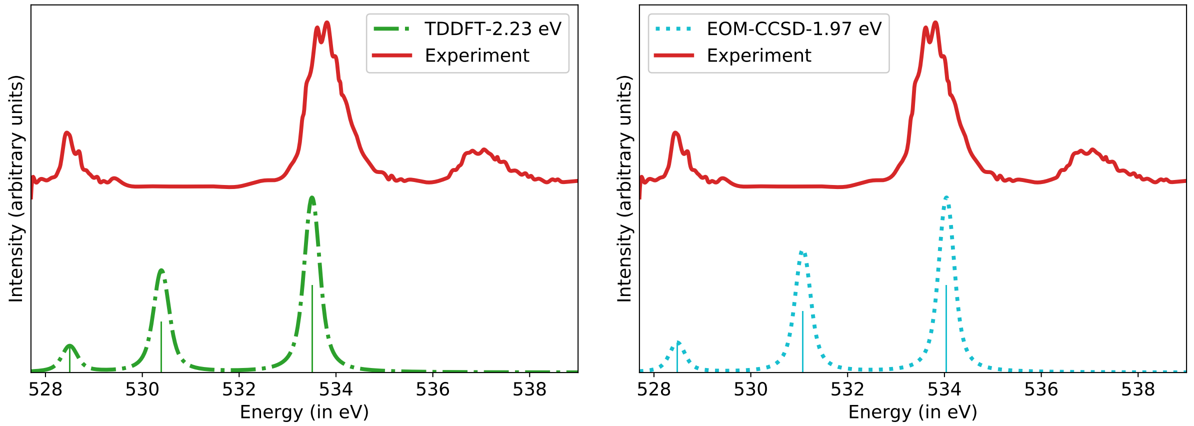
Fig 13.3 presents the orbital optimized SCAN spectrum (both recoupled and mixed), along with those from translated TDDFT and fc-CVS-EOM-CCSD. There are three peaks in all cases: the 1s→ SOMO excitation (lowest in energy) and the two doublets arising from 1s→LUMO excitations. We observe that the linear-response approaches yield a fairly poor picture. Both TD-SRC2-R1 (Fig 13.3c) and EOM-CCSD (Fig 13.3d) need to be redshifted by ~ 2 eV to align the 1s → SOMO peak with experiment (vs the orbital optimized DFT spectra, which needs no such translation). The translated spectra are nonetheless greatly compressed relative to experiment and the relative intensities of the two 1s→ LUMO peaks are incorrect. This is not merely a consequence of spin-contamination, as Fig 13.3b shows that SCAN using mixed configurations does better at reproducing the overall shape of the spectrum, despite having quartet contamination as well. Lack of orbital relaxation thus appears to be the critical factor that compromises the performance of TDDFT and EOM-CCSD for this system.

Fig 13.3b however also shows that SCAN with mixed configurations has too small a spacing between the two 1s→LUMO doublets (the two highest energy peaks). Fig 13.3a demonstrates that recoupling fixes this problem (and correctly reduces the intensity of the highest energy peak), yielding a spectrum is in decent agreement with experiment. The spacing between the two highest energy peaks remains somewhat small (2.8 eV) vs experiment (~3.4 eV but the unresolved broadness of the experimental second peak makes this



(a) Recoupled configurations/SCAN

(b) Mixed Configurations/SCAN



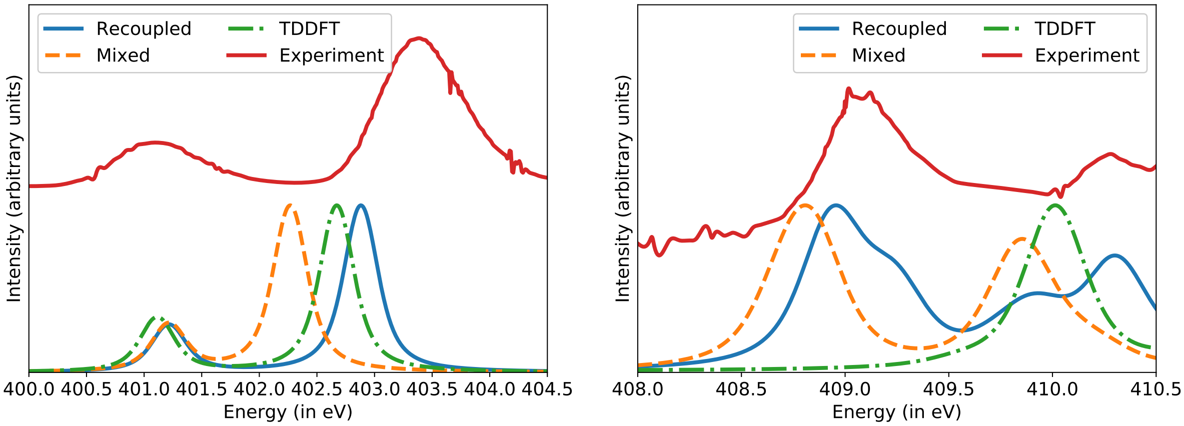
(c) TDDFT/SRC2-R1

(d) fc-CVS-EOM-CCSD [536]

Figure 13.3: Comparison of experimental O K-edge spectrum of CO^+ (obtained from Ref 560) with those computed with DFT/aug-cc-pCVTZ. Translated TDDFT and fc-CVS-EOM-CCSD [536] are also provided for comparison. A Voigt profile with a Gaussian standard deviation of 0.1 eV and Lorentzian $\gamma = 0.121$ eV was utilized for broadening the computed spectra. Bars are supplied to denote the location of the predicted excitation energies.

hard to pinpoint). Other DFT functionals similarly underestimate this splitting (to varying extents), while reproducing the general shape of the spectrum (as can be seen from the Supporting Information). Nonetheless, it is undeniable that the spectrum quality is greatly improved by recoupling. We also note that the NOCIS method [376, 477, 496] (which performs linear-response atop orbitals relaxed for the core-ionized state and is spin-pure in a manner analogous to XCIS [322]) yields spectra in excellent agreement with experiment (as shown in the Supporting Information), further demonstrating the utility of orbital relaxation and configuration recoupling, in an unambiguous, wavefunction based manner. At any rate, the qualitative failure of TDDFT and EOM-CCSD seems to argue for the use of methods with explicit orbital relaxation and configuration recoupling (like the scheme presented here or NOCIS) for the computation of core-level spectra of open-shell systems, irrespective of whether the computed spectra is translated or not.

13.3.4 N K-edge Spectrum of NO_2



(a) Valence excitations.

(b) Rydberg excitations.

Figure 13.4: Comparison of experimental N K-edge spectrum of NO_2 (obtained from Ref 564) with those computed with DFT/d-aug-cc-pCVTZ [70, 198, 437, 570] for both the valence (left) and Rydberg (right) regimes. The actual intensities of the Rydberg states are roughly an order of magnitude lower than that of the valence states, but have been magnified for easier comparison. The SRC2-R1 functional was employed for the TDDFT spectrum, while SCAN was utilized for both the recoupled and mixed configuration approaches. A Voigt profile with a Gaussian standard deviation of 0.1 eV and Lorentzian $\gamma = 0.121$ eV was utilized for broadening the computed spectra.

NO_2 is another rare open-shell system with a known experimental high resolution core-level spectrum [564], by virtue of being quite stable for a radical. It is isoelectronic with allyl, although the SOMO is not a π^* orbital (but rather resembles a N lone 'pair' orbital). The spectrum is nonetheless dominated by the transitions to the SOMO and the π^* LUMO levels. However, some Rydberg states have also been characterized, indicating that it could serve as an example to demonstrate whether our approach is balanced at predicting both valence and Rydberg excitations simultaneously.

Fig 13.4 compares the experimental spectrum at the N K-edge with those predicted via DFT (employing the doubly augmented d-aug-cc-pCVTZ basis to properly converge Rydberg states). The valence regime spectrum in Fig 13.4a shows that all methods get the qualitative form right, though the 1s to π^* LUMO transition is somewhat redshifted by all methods.

The success of TDDFT here stands in contrast to the failure observed for the valence regime of the allyl radical, although the different symmetry of the SOMO (lone 'pair' like vs π^*) may contribute to this. Recoupled SCAN performs better than mixed configuration SCAN for the second excitation by removing the quartet contribution to the energy. This blueshifts the 402.3 eV excitation energy predicted by the mixed configuration approach to 402.9 eV, which is much closer to the experimentally observed peak at 403.3 eV. This disagreement is not particularly small (and is in the opposite direction to the systematic overestimation exhibited by SCAN for excitations to the SOMO), but the recoupled DFT method gives best agreement with experiment.

The Rydberg regime depicted in Fig 13.4b however shows somewhat surprising behavior. It was tempting to believe that the weak coupling between the excited electron and the other unpaired electrons would lead to good performance by all methods. However, TDDFT absolutely fails to reproduce the spectrum in this regime, significantly blueshifting the experimental peak at 408.9 eV to 410.0 eV. On the other hand, the mixed configurations are quartet contaminated, and are thus slightly redshifted from their optimal location. Our recoupling protocol eliminates this problem, giving excellent agreement with experiment. It is also worth noting that the recoupled approach appears to predict the shape of the curve better than individual mixed configurations, indicating that the protocol described in Sec 13.2.3 was reasonably effective. This is however ultimately only one data point, and comparison against more high resolution experimental spectra would be useful in validating our observation. We therefore hope that spectra of more open-shell species in the Rydberg regime will be available in the near future. We note that high energy spectra for N_2^+ and CO^+ have been very recently reported [560, 562], but the Rydberg region appears to also contain a large number of doubly excited states with significant multiconfigurational character (involving more than three orbitals), that DFT based methods are unlikely to successfully model. This is less likely to be the case for neutral species.

13.4 Recommendations for Successful Calculations

The proposed protocol for recoupling mixed configurations appears to yield improved agreement with experiment relative to simply using the two individual mixed configurations that correspond to single excitations. Nonetheless, it entails individual optimization of four configurations per excitation ($|Q, M_{1,2,3}\rangle$), to get two doublet state energies. We subsequently recommend the following protocol for ensuring maximum agreement between these configurations and minimizing computational cost.

1. Optimize unrestricted KS ground state orbitals.
2. Use these orbitals as initial guesses to optimize RO orbitals for the ground state.

3. Using the RO ground state orbitals as the initial guess, optimize the RO orbitals for the core-ionized state via SGM. This decouples the relaxation of the core-hole from the rest of the computations.

4. Using the RO core-ionized orbitals as the initial guess, optimize RO orbitals corresponding to the desired quartet state with SGM. The core-ionized orbitals can thus be computed only once, and repeatedly utilized for multiple excitations. Furthermore, the unoccupied orbitals for the core ionized state are much more representative of the optimized orbitals for the excited electron, than canonical ground state orbitals.

5. Using the RO core-excited quartet orbitals as initial guesses, find the unrestricted orbitals for the quartet $|Q\rangle$ and mixed configurations $|M_{1,2,3}\rangle$ with SGM.

Steps 1-3 also apply for excitations to the SOMO level, followed by use of the RO core ionized orbitals to initialize the excited state optimization for the core to SOMO excited configuration. They also apply for computation of core-excitations in closed-shell species via ROKS. We believe that the RO energies themselves are not particularly useful for radicals, but the RO orbitals act as useful intermediates to prevent the alpha and beta spatial orbitals from differing prior to the last optimization step (step 5). The RO orbital space in fact is much more tightly constrained and SGM is faster at those optimizations in practice. Difficult convergence cases in general could also be addressed via converging to the same state with a different (ideally, cheaper) functional and using the resulting orbitals as initial guesses.

Three additional points regarding orbital optimized core-excitation calculations in general (for both closed and open-shell systems) are worth noting as well.

1. Use of a localized core-hole is absolutely critical for systems where there are symmetry equivalent atoms (like the terminal carbons of allyl). Delocalized core-holes lead to substantial delocalization error [64, 95] driven underestimation of energy, as shown in Ref 546. Localization of core orbitals can be achieved via explict localization, or via weak electric fields that break symmetry. The mixed basis strategy described in the next point also leads to core orbital localizing symmetry breaking.

2. It is absolutely essential to use at least a triple zeta level basis with split core functions (like cc-pCVTZ) at the local site of the core-excitation. The core-hole would otherwise not be able to adequately relax, and energies be systematically overestimated [546]. However, a smaller basis can be used for all other atoms, with cc-pVDZ being adequate in our experience [546] (though even smaller bases could potentially be fine). This mixed basis strategy helps bring down the computational cost considerably as well, as the overall computation cost is comparable to a double zeta basis DFT ground state calculation per iteration, though excited state orbital optimization does often require many more iterations than ground state computations.

3. Many core-excited states possess significant Rydberg character. A good description of these states necessitates the presence of diffuse functions in the basis, and even double

augmentation is sometimes necessary (such as the NO_2 spectrum presented in Fig 13.4, where singly augmented aug-cc-pCVTZ blueshifts the Rydberg peaks in Fig 13.4b by 0.2 eV). This is easily the most onerous basis set requirement for such calculations but is functionally unavoidable for any electronic structure method seeking a correct description of Rydberg states.

13.5 Conclusions

We have investigated orbital optimized density functional approaches to studying core-excitation spectra of open-shell systems, by employing the SGM approach for averting variational collapse. Lack of gas-phase experimental data proves to be a hindrance for assessing the performance of these methods, but existing data shows encouraging behavior. We firstly find that several density functionals like SCAN, TPSS, BLYP, B3LYP, cam-B3LYP and ωB97X-D3 can be employed to predict excitation energies corresponding to 1s to SOMO transitions in radicals, to RMSE at or below 0.3 eV. The 1s$\rightarrow$ SOMO transitions are however not very challenging excitations as they do not result in a change in the total number of unpaired electrons and thus can be well approximated by single Slater determinants.

Higher excitations entail breaking of electron pairs and thus are natively multiconfigurational. These states therefore cannot be described by single determinants, although somewhat reliable results can at times be obtained from symmetry broken mixed determinants in the limit of weak coupling between unpaired spins (analogous to how unrestricted HF/DFT being effective for single bond dissociations in closed-shell species). For more general accuracy, we present a CI inspired approach for self-consistently recoupling these single determinant mixed configurations with unpaired spins to yield approximately spin-pure results corresponding to multiconfigurational doublet states. The performance of this approach is compared against that of using unrecoupled mixed determinants alone and TDDFT for the core-level spectra of the allyl radical, CO^+ at the O K-edge and NO_2 at N K-edge. We also considered the performance of fc-CVS-EOM-CCSD for the allyl radical and the O K-edge of CO^+ . We find that the recoupling scheme leads to no degradation of performance and in fact consistently improves upon results obtained by merely using single mixed determinants (significantly so for the O K-edge of CO^+). It is nonetheless worth appreciating that unrecoupled determinants often yield fairly reasonable answers by themselves, especially relative to TDDFT /EOM-CCSD for the allyl radical and the O K-edge of CO^+. Our work therefore shows promise in using orbital optimized DFT approaches for predicting core-level spectra of radicals, where high accuracy can be obtained even from local functionals like SCAN, at low computational cost. Available evidence also appears to argue for recoupling mixed configurations, although this is roughly computationally twice as expensive (as four configurations need to be optimized as opposed to only two). The O K-edge of CO^+ also seems to suggest that our recoupling scheme somewhat underestimates doublet-doublet splitting in the strong coupling limit. More experimental spectra for open-shell systems (involving transitions to unoccupied valence orbitals) would however be immensely useful in fully characterizing the limitations of the recoupling approach. We consequently will continue to attempt to validate this approach via comparison to experiment as new data arises.

In future, we will also seek to develop approaches that optimize a single set of unrestricted orbitals for recoupling mixed configurations vs separately optimizing all four relevant states. This should reduce the computational cost of such calculations substantially, and enhance their utility. It would also be useful to generalize the recoupling approach to higher spin states like triplets, where there are more spins to recouple and a correspondingly larger number of coupling constants. Work along these directions is presently in progress.

13.6 Computational Methods

All calculations were performed with the Q-Chem 5.3 [191] package. Local exchange-correlation integrals were calculated over a radial grid with 99 points and an angular Lebedev grid with 590 points. Experimental geometries (from the NIST database [192]) were used whenever possible, with MP2 [19]/cc-pVTZ [198] optimized geometries being employed in their absence. The plots labeled ‘mixed’ only used the two mixed configurations corresponding to single excitations from the ground state, as the third configuration is technically a double excitation that would not usually be considered due to formally zero (and in practice, typically small) oscillator strength. All TDDFT calculations employed the Tamm-Dancoff Approximation [251, 313–315].

13.7 Appendix: Phase Estimation for Mixed Configurations

The phase convention chosen for $|M_{1,2,3}\rangle$ in Sec 13.2.2 ensures that the off-diagonal elements of the coupling matrix are $-J_{ij}$, where the couplings J_{ij} are given by:

$$J_{12} = \frac{E_{M_1} + E_{M_2} - E_Q - E_{M_3}}{2} \tag{13.4}$$

$$J_{13} = \frac{E_{M_1} + E_{M_3} - E_Q - E_{M_2}}{2} \tag{13.5}$$

$$J_{23} = \frac{E_{M_2} + E_{M_3} - E_Q - E_{M_1}}{2} \tag{13.6}$$

However, the determinants $|M_i'\rangle$ obtained from orbital optimization can differ from this ideal phase. Specifically, DFT can yield $|M_i'\rangle = p_i\,|M_i\rangle$ where $p_i = \pm 1$. This has no implication for the energies, but will affect properties like the transition dipole moment for which the relative phases of the configurations matter (as these properties depend on off-diagonal elements, and are computed from $|M_i'\rangle$ vs the idealized $|M_i\rangle$).

The easiest route for phase finding seems to be via the quartet state, which has the eigenvector $\left(\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}}\right)^T$ in the $|M_i\rangle$ basis. This state should formally have zero transition dipole moment and thus could be employed to compute relative phases.

Specifically, let $|M_i'\rangle$ have transition dipole moments $\vec{\mu}_i = \langle 0| \hat{\vec{\mu}} |M_i'\rangle$ against the ground state determinant $|0\rangle$. Without loss of generality, we can set the phase p_1 of $|M_1'\rangle$ to 1 (as only relative phases matter). Then the transition dipole moment of the ostensibly quartet state is $\vec{\mu}_Q = \frac{\vec{\mu}_1 + p_2\vec{\mu}_2 + p_3\vec{\mu}_3}{\sqrt{3}}$. Consequently, the signs of $p_{2,3}$ should be chosen to minimize this quantity. In practice, this protocol is often simplified on account of one of the three mixed determinants being a formal double excitation ($|M_2\rangle$ if the orbitals are ordered by energy), which would have typically have very low transition dipole moment (though generally not exactly zero on account of non-orthogonality between the ground and mixed determinant orbitals). The phase estimation problem here is thus often just finding whether p_3 (say) should be 1 or -1 for $\vec{\mu}_Q$ to be smallest.

In fact, this is essentially an internal consistency check for determining the impact of neglecting non-orthogonality between mixed determinants and the overall quality of the optimized orbitals, as this "quartet" transition dipole moment $\vec{\mu}_Q$ should be at least an order of magnitude smaller (and hopefully even less) than the largest transition dipole moment corresponding the two doublet states, after finding optimal phases. The oscillator strength scales as square of the transition dipole and thus any spurious "quartet" peak stemming from neglect of non-orthogonality etc. would be at least a hundred times weaker than the strongest doublet peak and thus the quality of the spectrum will be preserved.

As an example, let us consider the N 1s→ π^* transition in NO_2. The orbital optimized determinants $\left|M_{1,2,3}'\right\rangle$ we obtained had transition dipole moments (after ignoring terms smaller than 10^{-4}). :

$$\vec{\mu}_1 = -6.11 \times 10^{-2}\hat{x} \tag{13.7}$$

$$\vec{\mu}_2 = 0 \tag{13.8}$$

$$\vec{\mu}_3 = 5.98 \times 10^{-2}\hat{x} \tag{13.9}$$

$\vec{\mu}_Q$ is minimized if $p_3 = 1$, as then the dipoles will mostly cancel each other.

13.8 Supporting Information

Supporting information for this work can be accessed via the journal article. They include: additional spectra for the allyl radical and CO^+, raw data, geometries.

Chapter 14

Afterword to Electronic Excited States

And, by the incantation of this verse,
Scatter, as from an unextinguish'd hearth
Ashes and sparks, my words among mankind!
Be through my lips to unawaken'd earth
The trumpet of a prophecy! O Wind,
If Winter comes, can Spring be far behind?
Percy Bysshe Shelley. "Ode to the West Wind"

The four preceding chapters have reported on routes to apply DFT to excited states. Specifically, Chapter 10 was an in-depth investigation on how TDDFT style methods fail to model excited state bond dissociations, while Chapters 11-13 focus on OO-DFT for excited states, with an emphasis on core-level excitations. It therefore seems reasonable to take a comprehensive look at the presented work, as well as comment on the current state of modeling electronic excitations with OO-DFT. For a more detailed view, we encourage the reader to peruse Ref 366, which is our perspective article on OO-DFT.

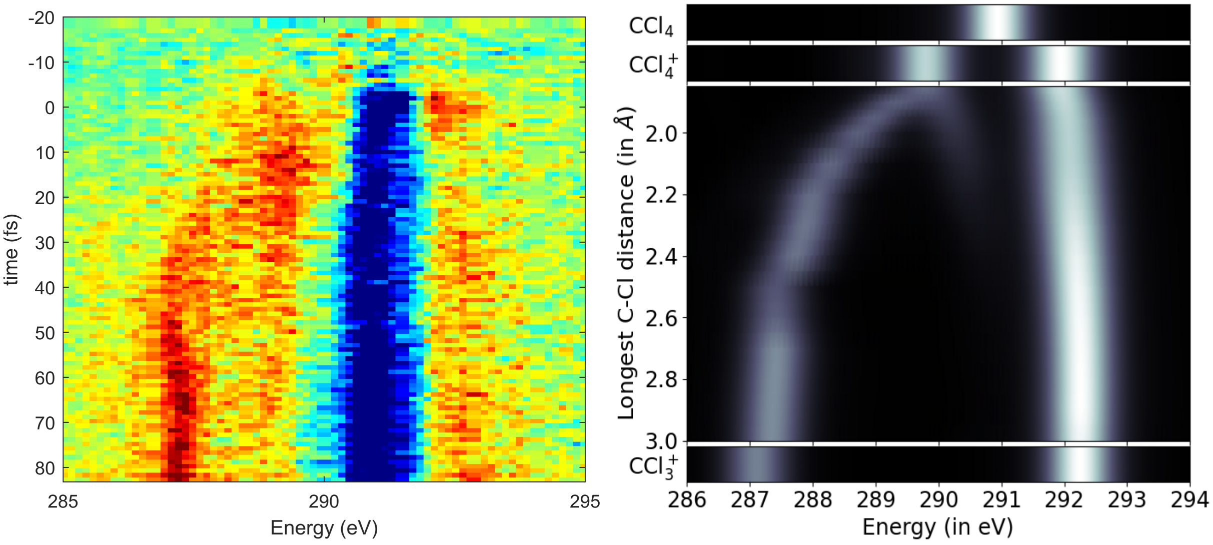
There has been considerable recent interest [344, 571–575] in developing algorithms free of variational collapse, to model excited states with OO-DFT methods. SGM is one of the 'earlier' members of the current generation of methods, and as such attracted some positive attention [572, 573, 576–578]. In turn, SGM itself was greatly inspired by wavefunction theory work done by the Neuscamman [351, 352] and Van Voorhis [353, 354] groups. As a direct minimization method that costs 2-3 times as much as routine ground state KS-DFT (per iteration), SGM is certainly one of the more computationally demanding excited state optimizers. However, being a direct minimization based algorithm, SGM is likely more robust against variational collapse than repeated diagonalization based methods or saddle point solvers. It is certainly possible that no excited state optimization algorithm would become ubiquitous the way DIIS [342] has emerged for the ground state. Even if SGM only becomes an 'optimizer of last resort' for difficult problems in the future, it would have still served an important purpose in stimulating OO-DFT research and demonstrating the possibility of reliable, variational collapse free OO-DFT from a direct minimization perspective. Ultimately, the success of OO-DFT methods should not be tied to a single optimizer, but rather what the class of such algorithms can collectively achieve.

And they can achieve a great deal! OO-DFT is effective at computing CT excited states where standard TDDFT fails, such as in organic light emitting diode materials [265, 579]. This indicates that OO-DFT approaches can also prove quite useful in modeling photoactive materials with large levels of charge separation, like photovoltaics or photosynthetic systems. OO-DFT can also effectively access doubly excited electronic states that are absent in TDDFT, as shown in Chapter 11. OO-DFT's ability to access traditionally challenging states also does not come at the cost of modeling 'simple' singly excited states [441], for which it appears to be marginally more accurate than TDDFT [319].

The efficacy of OO-DFT in predicting core-level spectra merits particular attention. OO-DFT/SCAN delivers semiquantitative accuracy ($\sim 0.2 - 0.3$ eV root mean square error) for the K-edges of light elements C, N, O, and F against experiment (~ 0.1 eV uncertainty), as reported in Chapters 12 and 13. This level of accuracy for core excited states has not been surpassed by any quantum chemistry method with better than $O(N^7)$ scaling, to the best of our knowledge. The success of OO-DFT/SCAN for this application is almost entirely the consequence of orbital optimization, as OO-DFT with the completely correlation free HF functional has rather low error [366] (0.6 eV root mean squared error over the dataset in Chapter 12) relative to CIS/TDDFT deviations from experiment (~ 10 eV). Nonetheless, the SCAN DFA plays a nonnegligible role in lowering the error further to ~ 0.2 eV, to the point that it becomes comparable to the typical experimental uncertainty of ~ 0.1 eV.

In terms of developments beyond what has been reported in Chapters 12 and 13, we have recently included relativistic effects into OO-DFT through the spin-free exact two component one electron model [580]. This permits application of OO-DFT to model K-edges of heavier elements [581], going up to Cr (~ 6000 eV, with ~ 40 eV relativistic contribution). A fully *ab initio*, relativistic treatment of spin-orbit effects is presently under development for studying L-edge spectra and beyond. Efforts are also underway to determine if both the particle and hole orbitals for a core-level excited state need to be fully optimized, or whether it is sufficient to merely account for the relaxation of the core-hole and then perform linear-response to determine particle levels [582]. Such a hybrid protocol is considerably more computationally efficient for core excited states, by virtue of being core-hole specific and not state specific. This approach can also be used to cheaply filter states with significant absorbance, so that only such states are fully orbital optimized and no extra effort is wasted on weakly absorbing states.

OO-DFT can also be used to interpret new experiments, not merely reproduce preexisting results. We are presently investigating the femtosecond resolved dynamics of strong field ionized CCl_4, in collaboration with experimentalists from the Leone and Neumark groups. Experimental C K-edge spectra and OO-DFT calculations collectively describe the cleavage of a covalent bond, resulting in CCl_3^+ and atomic Cl. A comparison between experiment and OO-DFT results is provided at Fig 14.1. Comparison suggests that CCl_4^+ splits into CCl_3^+ and Cl starting around 40 fs. In principle, OO-DFT calculations are therefore capable of



(a) Experiment.

(b) Theory.

Figure 14.1: C K-edge X-ray absorption spectra for dissociation of $CCl_4{}^+$. The experimental time-resolved transient spectrum is given on the left with the negative signal (blue) at $\sim$ 291 eV corresponding to depletion of neutral CCl_4 while positive signal (red) corresponds to formation of new species. In particular, positive signal arises at $\sim$289.5 eV and $\sim$ 292 eV at short times, with the $\sim$289.5 eV signal evolving to $\sim$ 287 eV over time. The OO-DFT/SCAN spectrum on the right is neither time-resolved nor transient, but merely shows C-K edge absorption for different species (white signal). Neutral CCl_4 is depicted on the top (absorbing at $\sim$ 291 eV), followed by the C_{2v} geometry of $CCl_4{}^+$ (which is the only minimum with four covalent C-Cl bonds) that absorbs at $\sim$289.5 eV and $\sim$ 292 eV. The absorption of $CCl_4{}^+$ configurations where the longest C-Cl bond is constrained to specific values between 1.85-3 Å is also shown, depicting evolution of signal on the route to dissociation. The $\sim$289.5 eV absorption signal redshifts with increasing C-Cl stretch, reaching $\sim$ 287 eV around $\sim$ 3 Å, which is very close to absorption of isolated $CCl_3{}^+$. This indicates near complete covalent bond cleavage leads to the $\sim$ 287 eV experimental signal.
This figure utilizes material from Ross, A.D; **Hait, D.**; Scutelnic, V; Haugen, E.A.; Ridente, E.; Balkew, M.; Neumark, D.M.; Head-Gordon, M.; Leone, S.R; "Jahn-Teller Distortion and Dissociation of $CCl_4{}^+$ by Transient X-ray Spectroscopy Simultaneously at the Carbon K- and Chlorine L-Edge" *(in preparation)*, with permission from all authors.

directly simulating experimental observables. OO-DFT can thus be used to interpret other time-resolved experimental spectra involving dynamics of excited states, and we are pursuing a number of such applications at present.

OO-DFT also has a symbiotic relationship with wavefunction theory, with developments on one side influencing the other. Some very interesting recent work [571, 583] on the energy landscape for excited states is particularly instructive about understanding such states. Excited state specific MP2 [441] and coupled cluster approaches [584, 585] are also being developed. State specific active space methods have also been proposed for strongly correlated systems [586]. It therefore seems likely that parallel developments in the general area of excited state specific quantum chemistry are going to continue in the near future from both wavefunction and DFT perspectives. One particularly interesting area of synergy would be the use of double hybrid functionals for excited state OO-DFT, which has not seen much investigation at present but has the possibility to yield more accurate results.

In some regards, the successes of the present state of OO-DFT are astounding as good results are obtained with ground state DFAs that were designed without any excited state information. Future functional development that utilizes excited state information explicitly has the potential to yield models with greater effectiveness for OO-DFT applications. Such DFAs would also be better approximations to the exact functional as they would utilize information about multiple extrema of the functional, not just the ground state minima [117]. Indeed, functionals trained with excited state information may even improve ground state predictions through improvements in transferability, although this is not certain.

In summary, OO-DFT approaches are almost certainly going to play key roles in efficiently modeling optical properties of materials and experimental spectra of larger molecules. The broader community therefore should embrace such state specific methods wholeheartedly for problems that the widely used linear-response paradigm is challenged by. Many of the problems historically faced by such methods have been resolved and further research can only improve the situation. It would therefore be interesting to see how the general area evolves over the coming years.

Bibliography

(1) Feynman, R.; Leighton, R.; Sands, M., *The Feynman Lectures on Physics*; The Feynman Lectures on Physics v. 1; Addison-Wesley Publishing Company: 1965.

(2) Dirac, P. A. M. Quantum mechanics of many-electron systems. *Proc. R. Soc. A* **1929**, *123*, 714–733.

(3) Born, M.; Oppenheimer, R. Zur quantentheorie der molekeln. *Ann. der Phys.* **1927**, *389*, 457–484.

(4) Tully, J. C. Perspective on "zur quantentheorie der molekeln". *Theor. Chem. Acc.* **2000**, *103*, 173–176.

(5) Sakurai, J. J., *Modern quantum mechanics; rev. ed.* Addison-Wesley: Reading, MA, 1994.

(6) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: molecular dynamics through a conical intersection. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127–158.

(7) McQuarrie, D. A., *Quantum chemistry*; University Science Books: 2008.

(8) Rydberg, J. XXXIV. On the structure of the line-spectra of the chemical elements. *Lond. Edinb. Dubl. Phil. Mag.* **1890**, *29*, 331–337.

(9) Gerlach, W.; Stern, O. Der experimentelle nachweis der richtungsquantelung im magnetfeld. *Z. Phys.* **1922**, *9*, 349–352.

(10) Fermi, E. Sulla quantizzazione del gas perfetto monoatomico. *Rend. Lincei* **1926**, *145*.

(11) Dirac, P. A. M. On the theory of quantum mechanics. *Proc. R. Soc. A* **1926**, *112*, 661–677.

(12) Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Z. Phys.* **1925**, *31*, 765–783.

(13) Slater, J. C. The theory of complex spectra. *Phys. Rev.* **1929**, *34*, 1293.

(14) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, New York, 1996, pp 286–296.

(15) Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XVIII. Constraints and stability in Hartree–Fock theory. *J. Chem. Phys.* **1977**, *66*, 3045–3050.

(16) Hartree, D. R. In *Math. Proc. Camb. Philos. Soc.* 1928; Vol. 24, pp 89–110.

(17) Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.* **1930**, *61*, 126–148.

(18) Slater, J. C. Note on Hartree's method. *Phys. Rev.* **1930**, *35*, 210.

(19) Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 509–530.

(20) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

(21) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(22) Boys, S. F.; Handy, N. C. A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy. *Proc. R. Soc. Lond. A* **1969**, *310*, 63–78.

(23) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(24) Vogiatzis, K. D.; Ma, D.; Olsen, J.; Gagliardi, L.; De Jong, W. A. Pushing configuration-interaction to the limit: Towards massively parallel MCSCF calculations. *J. Chem. Phys.* **2017**, *147*, 184111.

(25) Eriksen, J. J. et al. The ground state electronic energy of benzene. *J. Phys. Chem. Lett.* **2020**, *11*, 8922–8929.

(26) Löwdin, P.-O. Correlation Problem in Many-Electron Quantum Mechanics I. Review of Different Approaches and Discussion of Some Current Ideas. *Adv. Chem. Phys.* **1958**, 207–322.

(27) Garrod, C.; Percus, J. K. Reduction of the N-particle variational problem. *J. Math. Phys.* **1964**, *5*, 1756–1776.

(28) Coleman, A. J. Structure of fermion density matrices. *Rev. Mod. Phys.* **1963**, *35*, 668.

(29) Mazziotti, D. A. Variational reduced-density-matrix method using three-particle N-representability conditions with application to many-electron molecules. *Phy. Rev. A* **2006**, *74*, 032501.

(30) Li, R. R.; Liebenthal, M. D.; DePrince III, A. E. Challenges for variational reduced-density-matrix theory with three-particle N-representability conditions. *J. Chem. Phys* **2021**, *155*, 174110.

(31) Zhao, Z.; Braams, B. J.; Fukuda, M.; Overton, M. L.; Percus, J. K. The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions. *J. Chem. Phys* **2004**, *120*, 2095–2104.

(32) Erdahl, R. M. Representability. *Int. J. Quant. Chem.* **1978**, *13*, 697–718.

(33) Fosso-Tande, J.; Nguyen, T.-S.; Gidofalvi, G.; DePrince III, A. E. Large-scale variational two-electron reduced-density-matrix-driven complete active space self-consistent field methods. *J. Chem. Theory Comput.* **2016**, *12*, 2260–2271.

(34) Stair, N. H.; Evangelista, F. A. Exploring Hilbert space on a budget: Novel benchmark set and performance metric for testing electronic structure methods in the regime of strong correlation. *J. Chem. Phys* **2020**, *153*, 104108.

(35) Thomas, L. H. In *Math. Proc. Camb. Philos. Soc.* 1927; Vol. 23, pp 542–548.

(36) Fermi, E. Un metodo statistico per la determinazione di alcune priorieta dell'atome. *Rend. Accad. Naz. Lincei* **1927**, *6*, 32.

(37) Dirac, P. A. In *Math. Proc. Camb. Philos. Soc.* 1930; Vol. 26, pp 376–385.

(38) Wang, W.-P.; Parr, R. G.; Murphy, D. R.; Henderson, G. A. Gradient expansion of the atomic kinetic energy functional. *Chem. Phys. Lett.* **1976**, *43*, 409–412.

(39) Weizsäcker, C. v. Zur theorie der kernmassen. *Z. Phys.* **1935**, *96*, 431–458.

(40) Teller, E. On the stability of molecules in the Thomas-Fermi theory. *Rev. Mod. Phys.* **1962**, *34*, 627.

(41) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864.

(42) Gilbert, T. L. Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B* **1975**, *12*, 2111.

(43) Löwdin, P.-O. Twenty-five years of Sanibel symposia: A brief historic and scientific survey. *Int. J. Quantum Chem* **1986**, *28*, 19–37.

(44) Handy, N. C. In *Quantum mechanical simulation methods for studying biological systems*; Springer: 1996, pp 1–35.

(45) Kato, T. On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.* **1957**, *10*, 151–177.

(46) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6062–6065.

(47) Harriman, J. E. Orthonormal orbitals for the representation of an arbitrary density. *Phys. Rev. A* **1981**, *24*, 680.

(48) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133.

(49) Schipper, P.; Gritsenko, O.; Baerends, E. One-determinantal pure state versus ensemble Kohn-Sham solutions in the case of strong electron correlation: CH2 and C2. *Theor. Chem. Acc.* **1998**, *99*, 329–343.

(50) Mayer, I.; Pápai, I.; Bakó, I.; Nagy, A. Conceptual Problem with Calculating Electron Densities in Finite Basis Density Functional Theory. *J. Chem. Theory Comput.* **2017**, *13*, 3961–3963.

(51) Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **1999**, *71*, 1253.

(52) Pines, D. A collective description of electron interactions: IV. Electron interaction in metals. *Phys. Rev.* **1953**, *92*, 626.

(53) Ceperley, D. M.; Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **1980**, *45*, 566.

(54) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(55) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(56) Karton, A.; Daon, S.; Martin, J. M. W4-11: A high-confidence benchmark dataset for computational thermochemistry derived from first-principles W4 data. *Chem. Phys. Lett.* **2011**, *510*, 165–178.

(57) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(58) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(59) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(60) Perdew, J. P.; Constantin, L. A. Laplacian-level density functionals for the kinetic energy density and exchange-correlation energy. *Phys. Rev. B* **2007**, *75*, 155109.

(61) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(62) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(63) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.

(64) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz Jr, J. L. Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.* **1982**, *49*, 1691.

(65) Jensen, F. Polarization consistent basis sets: Principles. *J. Chem. Phys.* **2001**, *115*, 9113–9125.

(66) Jensen, F. Polarization consistent basis sets. II. Estimating the Kohn–Sham basis set limit. *J. Chem. Phys.* **2002**, *116*, 7372–7379.

(67) Jensen, F. Polarization consistent basis sets. III. The importance of diffuse functions. *J. Chem. Phys.* **2002**, *117*, 9234–9240.

(68) Jensen, F.; Helgaker, T. Polarization consistent basis sets. V. The elements Si–Cl. *J. Chem. Phys.* **2004**, *121*, 3463–3470.

(69) Jensen, F. Polarization Consistent Basis Sets. 4: The Elements He, Li, Be, B, Ne, Na, Mg, Al, and Ar. *J. Phys. Chem. A* **2007**, *111*, 11198–11204.

(70) Woon, D. E.; Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(71) Yang, W.; Zhang, Y.; Ayers, P. W. Degenerate ground states and a fractional number of electrons in density and reduced density matrix functional theory. *Phys. Rev. Lett.* **2000**, *84*, 5172.

(72) Ayers, P. W. The dependence on and continuity of the energy and other molecular properties with respect to the number of electrons. *J. Mat. Chem.* **2008**, *43*, 285–303.

(73) Ayers, P. W.; Levy, M. Levy Constrained Search in Fock Space: An Alternative Approach to Noninteger Electron Number. *Acta Phys. Chim. Sin.* **2017**, *34*, 625–630.

(74) Perdew, J. P. In *Density Functional Methods in Physics*; Springer: 1985, pp 265–308.

(75) Zhang, Y.; Yang, W. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. *J. Chem. Phys.* **1998**, *109*, 2604–2608.

(76) Gräfenstein, J.; Kraka, E.; Cremer, D. The impact of the self-interaction error on the density functional theory description of dissociating radical cations: Ionic and covalent dissociation limits. *J. Chem. Phys.* **2004**, *120*, 524–539.

(77) Vydrov, O. A.; Scuseria, G. E.; Perdew, J. P. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.* **2007**, *126*, 154109.

(78) Dutoi, A. D.; Head-Gordon, M. Self-interaction error of local density functionals for alkali-halide dissociation. *Chem. Phys. Lett.* **2006**, *422*, 230–233.

(79) Ruzsinszky, A.; Perdew, J.; Csonka, G.; Vydrov, O.; Scuseria, G. Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals. *J. Chem. Phys.* **2006**, *125*, 194112.

(80) Hait, D.; Head-Gordon, M. How accurate is density functional theory at predicting dipole moments? An assessment using a new database of 200 benchmark values. *J. Chem. Theory Comput.* **2018**, *14*, 1969–1981.

(81) Hait, D.; Head-Gordon, M. Communication: xDH double hybrid functionals can be qualitatively incorrect for non-equilibrium geometries: Dipole moment inversion and barriers to radical-radical association using XYG3 and XYGJ-OS. *J. Chem. Phys.* **2018**, *148*, 171102.

(82) Patchkovskii, S.; Ziegler, T. Improving "difficult" reaction barriers with self-interaction corrected density functional theory. *J. Chem. Phys.* **2002**, *116*, 7806–7813.

(83) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.* **2006**, *125*, 201102.

(84) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. Adiabatic connection for kinetics. *J. Phys. Chem. A* **2000**, *104*, 4811–4815.

(85) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Fractional charge perspective on the band gap in density-functional theory. *Phys. Rev. B* **2008**, *77*, 115123.

(86) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Phys. Rev. Lett.* **2008**, *100*, 146401.

(87) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E. Density functionals that are one-and two-are not always many-electron self-interaction-free, as shown for H_2^+, He_2^+, Li H^+, and Ne_2^+. *J. Chem. Phys.* **2007**, *126*, 104102.

(88) Li, C.; Yang, W. On the piecewise convex or concave nature of ground state energy as a function of fractional number of electrons for approximate density functionals. *J. Chem. Phys.* **2017**, *146*, 074107.

(89) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(90) Becke, A. D. Perspective: Fifty years of density-functional theory in Chem. Phys. *J. Chem. Phys.* **2014**, *140*, 18A301.

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

(92) Boese, A. D.; Martin, J. M. Development of density functionals for thermochemical kinetics. *J. Chem. Phys.* **2004**, *121*, 3405–3416.

(93) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(94) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(95) Hait, D.; Head-Gordon, M. Delocalization Errors in Density Functional Theory are Essentially Quadratic in Fractional Occupation Number. *J. Phys. Chem. Lett.* **2018**, *9*, 6280–6288.

(96) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(97) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(98) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ω B97M (2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.

(99) Perdew, J. P.; Schmidt, K. In *AIP Conference Proceedings*, 2001; Vol. 577, pp 1–20.

(100) Wang, H. Formation of nascent soot and other condensed-phase materials in flames. *Proc. Combust. Inst.* **2011**, *33*, 41–67.

(101) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(102) Grimme, S. On the Importance of Electron Correlation Effects for the π-π Interactions in Cyclophanes. *Chem. Eur. J.* **2004**, *10*, 3423–3429.

(103) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comp. Chem.* **2004**, *25*, 1463–1473.

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

(105) Rezác, J.; Riley, K. E.; Hobza, P. Extensions of the S66 data set: more accurate interaction energies and angular-displaced nonequilibrium geometries. *J. Chem. Theory Comput.* **2011**, *7*, 3466–3470.

(106) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.

(107) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

(108) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(109) Cohen, A.; Mori-Sánchez, P.; Yang, W. Fractional spins and static correlation error in density functional theory. *J. Chem. Phys.* **2008**, *129*, 121104–121104.

(110) Becke, A. D. Density functionals for static, dynamical, and strong correlation. *J. Chem. Phys.* **2013**, *138*, 074109.

(111) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(112) Najibi, A.; Goerigk, L. The Nonlocal Kernel in van der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ωB97M-V Approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738.

(113) Santra, G.; Sylvetsky, N.; Martin, J. M. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129–5143.

(114) Kirkpatrick, J. et al. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science* **2021**, *374*, 1385–1389.

(115) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **2017**, *355*, 49–52.

(116) Hammes-Schiffer, S. A conundrum for density functional theory. *Science* **2017**, *355*, 28–29.

(117) Perdew, J. P.; Levy, M. Extrema of the density functional for the energy: Excited states from the ground-state theory. *Phys. Rev. B* **1985**, *31*, 6264.

(118) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *J. Chem. Phys.* **2001**, *115*, 9233–9242.

(119) Boese, A. D.; Doltsinis, N. L.; Handy, N. C.; Sprik, M. New generalized gradient approximation functionals. *J. Chem. Phys.* **2000**, *112*, 1670–1678.

(120) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(121) Zhao, Y.; Truhlar, D. G. Density functional for spectroscopy: no long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

(122) Zhao, Y.; Truhlar, D. G. Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1849–1868.

(123) Peverati, R.; Truhlar, D. G. Improving the accuracy of hybrid meta-GGA density functionals by range separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.

(124) Peverati, R.; Truhlar, D. G. M11-L: A local density functional that provides improved accuracy for electronic structure calculations in chemistry and physics. *J. Phys. Chem. Lett.* **2011**, *3*, 117–124.

(125) Peverati, R.; Truhlar, D. G. An improved and broadly accurate local approximation to the exchange–correlation density functional: The MN12-L functional for electronic structure calculations in chemistry and physics. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13171–13174.

(126) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.* **2016**, *7*, 5032–5051.

(127) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A new local exchange-correlation functional for Kohn–Sham density functional theory with broad accuracy for atoms, molecules, and solids. *J. Chem. Theory Comput.* **2016**, *12*, 1280–1293.

(128) Graziano, G. Quantum chemistry: DFT's midlife crisis. *Nat. Rev. Chem.* **2017**, *1*, 0019.

(129) Kepp, K. P. Comment on "Density functional theory is straying from the path toward the exact functional". *Science* **2017**, *356*, 496–496.

(130) Korth, M. Density Functional Theory: Not Quite the Right Answer for the Right Reason Yet. *Angew. Chem. Int. Ed.* **2017**, *56*, 5396–5398.

(131) Brorsen, K. R.; Yang, Y.; Pak, M. V.; Hammes-Schiffer, S. Is the Accuracy of Density Functional Theory for Atomization Energies and Densities in Bonding Regions Correlated? *J. Phys. Chem. Lett.* **2017**, *8*, 2076–2081.

(132) Kepp, K. P. Density Functional Theory is Not Straying from the Path toward the Exact Functional. *arXiv preprint arXiv:1702.00813* **2017**.

(133) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Response to Comment on "Density functional theory is straying from the path toward the exact functional". *Science* **2017**, *356*, 496–496.

(134) Gould, T. What makes a density functional approximation good? Insights from the left Fukui function. *J. Chem. Theory Comput.* **2017**, *13*, 2373–2377.

(135) Ranasinghe, D. S.; Perera, A.; Bartlett, R. J. A note on the accuracy of KS-DFT densities. *J. Chem. Phys.* **2017**, *147*, 204103.

(136) Kullback, S.; Leibler, R. A. On information and sufficiency. *Ann. Math. Stats.* **1951**, *22*, 79–86.

(137) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897.

(138) Dirac, P. A. In *Proc. R. Soc. A*, 1931; Vol. 133, pp 60–72.

(139) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.

(140) Boese, A. D.; Handy, N. C. A new parametrization of exchange–correlation generalized gradient approximation functionals. *J. Chem. Phys.* **2001**, *114*, 5497–5503.

(141) Verma, P.; Truhlar, D. G. Can Kohn–Sham density functional theory predict accurate charge distributions for both single-reference and multi-reference molecules? *Phys. Chem. Chem. Phys.* **2017**, *19*, 12898–12912.

(142) Wang, Y.; Wang, X.; Truhlar, D. G.; He, X. How Well Can the M06 Suite of Functionals Describe the Electron Densities of Ne, Ne6+, and Ne8+? *J. Chem. Theory Comput.* **2017**, *13*, 6068–6077.

(143) Kepp, K. P. Energy vs. density on paths toward more exact density functionals. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7538–7548.

(144) Vanommeslaeghe, K. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(145) Ponder, J. W.; Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27–85.

(146) Ponder, J. W. et al. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.

(147) Hickey, A. L.; Rowley, C. N. Benchmarking quantum chemical methods for the calculation of molecular dipole moments and polarizabilities. *J. Phys. Chem. A* **2014**, *118*, 3678–3687.

(148) Puzzarini, C. Ab initio characterization of XH 3 (X= N, P). Part II. Electric, magnetic and spectroscopic properties of ammonia and phosphine. *Theor. Chem. Acc.* **2008**, *121*, 1–10.

(149) Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **1992**, *45*, 13244.

(150) Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **1981**, *23*, 5048.

(151) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.

(152) Zhang, Y.; Yang, W. Comment on “Generalized gradient approximation made simple". *Phys. Rev. Lett.* **1998**, *80*, 890.

(153) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. Workhorse semilocal density functional for condensed matter physics and quantum chemistry. *Phys. Rev. Lett.* **2009**, *103*, 026403.

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

(155) Csonka, G. I.; Perdew, J. P.; Ruzsinszky, A. Global hybrid functionals: a look at the engine under the hood. *J. Chem. Theory Comput.* **2010**, *6*, 3688–3703.

(156) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822.

(157) Perdew, J. P. et al. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **1992**, *46*, 6671.

(158) Yu, H. S.; Zhang, W.; Verma, P.; He, X.; Truhlar, D. G. Nonseparable exchange-correlation functional for molecules, including homogeneous catalysis involving transition metals. *Phys. Chem. Chem. Phys.* **2015**, *17*, 12146–12160.

(159) Peverati, R.; Truhlar, D. G. Exchange–correlation functional with good accuracy for both structural and energetic properties while depending only on the density and its gradient. *J. Chem. Theory Comput.* **2012**, *8*, 2310–2319.

(160) Zhao, Y.; Truhlar, D. G. Construction of a generalized gradient approximation by restoring the density-gradient expansion and enforcing a tight Lieb–Oxford bound. *J. Chem. Phys.* **2008**, *128*, 184109.

(161) Peverati, R.; Zhao, Y.; Truhlar, D. G. Generalized gradient approximation that recovers the second-order density-gradient expansion with optimized across-the-board performance. *J. Phys. Chem. Lett.* **2011**, *2*, 1991–1997.

(162) Peverati, R.; Truhlar, D. G. Communication: A global hybrid generalized gradient approximation to the exchange-correlation functional that satisfies the second-order density-gradient constraint and has broad applicability in chemistry. *J. Chem. Phys.* **2011**, *135*, 191102.

(163) Peverati, R.; Truhlar, D. G. Screened-exchange density functionals with broad accuracy for chemistry and solid-state physics. *Phys. Chem. Chem. Phys.* **2012**, *14*, 16187–16191.

(164) Mardirossian, N.; Ruiz Pestana, L.; Womack, J. C.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Use of the rVV10 Nonlocal Correlation Functional in the B97M-V Density Functional: Defining B97M-rV and Related Functionals. *J. Phys. Chem. Lett.* **2016**, *8*, 35–40.

(165) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(166) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

(167) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.* **2012**, *9*, 263–272.

(168) Lin, Y.-S.; Tsai, C.-W.; Li, G.-D.; Chai, J.-D. Long-range corrected hybrid meta-generalized-gradient approximations with dispersion corrections. *J. Chem. Phys.* **2012**, *136*, 154109.

(169) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *J. Chem. Phys.* **2005**, *123*, 161103.

(170) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.

(171) Sherrill, C. D.; Lee, M. S.; Head-Gordon, M. On the performance of density functional theory for symmetry-breaking problems. *Chem. Phys. Lett.* **1999**, *302*, 425–430.

(172) Sun, J.; Haunschild, R.; Xiao, B.; Bulik, I. W.; Scuseria, G. E.; Perdew, J. P. Semilocal and hybrid meta-generalized gradient approximations based on the understanding of the kinetic-energy-density dependence. *J. Chem. Phys.* **2013**, *138*, 044113.

(173) Sun, J.; Perdew, J. P.; Ruzsinszky, A. Semilocal density functional obeying a strongly tightened bound for exchange. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 685–689.

(174) Hui, K.; Chai, J.-D. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *J. Chem. Phys.* **2016**, *144*, 044114.

(175) Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(176) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.

(177) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.* **2006**, *125*, 224106.

(178) Rohrdanz, M. A.; Herbert, J. M. Simultaneous benchmarking of ground-and excited-state properties with long-range-corrected density functional theory. *J. Chem. Phys.* **2008**, *129*, 034107.

(179) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.* **2009**, *130*, 054112.

(180) Zhao, Y.; Truhlar, D. G. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.

(181) Boese, A. D.; Handy, N. C. New exchange-correlation density functionals: The role of the kinetic-energy density. *J. Chem. Phys.* **2002**, *116*, 9559–9569.

(182) Wellendorff, J.; Lundgaard, K. T.; Jacobsen, K. W.; Bligaard, T. mBEEF: An accurate semi-local Bayesian error estimation density functional. *J. Chem. Phys.* **2014**, *140*, 144107.

(183) Wang, Y.; Jin, X.; Haoyu, S. Y.; Truhlar, D. G.; He, X. Revised M06-L functional for improved accuracy on chemical reaction barrier heights, noncovalent interactions, and solid-state physics. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 8487–8492.

(184) Verma, P.; Bartlett, R. J. Increasing the applicability of density functional theory. IV. Consequences of ionization-potential improved exchange-correlation potentials. *J. Chem. Phys.* **2014**, *140*, 18A534.

(185) Jin, Y.; Bartlett, R. J. The QTP family of consistent functionals and potentials in Kohn-Sham density functional theory. *J. Chem. Phys.* **2016**, *145*, 034107.

(186) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Development of exchange-correlation functionals with minimal many-electron self-interaction error. *J. Chem. Phys.* **2007**, *126*, 191109.

(187) Chai, J.-D.; Head-Gordon, M. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.* **2009**, *131*, 174105.

(188) Zhang, Y.; Xu, X.; Goddard, W. A. Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 4963–4968.

(189) Zhang, I. Y.; Xu, X.; Jung, Y.; Goddard, W. A. A fast doubly hybrid density functional method close to chemical accuracy using a local opposite spin ansatz. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 19896–19900.

(190) Karton, A.; Tarnopolsky, A.; Lamére, J.-F.; Schatz, G. C.; Martin, J. M. Highly accurate first-principles benchmark data sets for the parametrization and validation of density functional and other approximate methods. Derivation of a robust, generally applicable, double-hybrid functional for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.

(191) Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

(192) Johnson III, R. D. NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 18. October 2016. *http://cccbdb.nist.gov/* **2015**.

(193) Takeo, H.; Sugie, M.; Matsumura, C. Microwave Spectroscopic Detection of Fluoroborane, BH2F. *J. Mol. Spectrosc.* **1993**, *158*, 201–207.

(194) Kawashima, Y.; Takeo, H.; Sugie, M.; Matsumura, C.; Hirota, E. Microwave spectrum of chloroborane, BH2Cl. *J. Chem. Phys.* **1993**, *99*, 820–826.

(195) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(196) Zhao, Y.; Truhlar, D. G. Benchmark databases for nonbonded interactions and their use to test density functional theory. *J. Chem. Theory Comput.* **2005**, *1*, 415–432.

(197) Tentscher, P. R.; Arey, J. S. Binding in radical-solvent binary complexes: Benchmark energies and performance of approximate methods. *J. Chem. Theory Comput.* **2013**, *9*, 1568–1579.

(198) Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(199) Peterson, K. A.; Dunning Jr, T. H. Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548–10560.

(200) Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theor. Chem. Acc.* **2011**, *128*, 69–82.

(201) Gill, P. M.; Johnson, B. G.; Pople, J. A. A standard grid for density functional calculations. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(202) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P. Basis-set convergence of the molecular electric dipole moment. *J. Chem. Phys.* **1999**, *111*, 4424–4430.

(203) Kurlancheek, W.; Head-Gordon, M. ller-Plesset perturbation theory and related double-hybrid density functional theory. *Mol. Phys.* **2009**, *107*, 1223–1232.

(204) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(205) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(206) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(207) Gwaltney, S. R.; Head-Gordon, M. A second-order perturbative correction to the coupled-cluster singles and doubles method: CCSD (2). *J. Chem. Phys.* **2001**, *115*, 2014–2021.

(208) Goerigk, L.; Grimme, S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals: Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2010**, *7*, 291–309.

(209) Kozuch, S.; Martin, J. M. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Comput. Chem.* **2013**, *34*, 2327–2344.

(210) Graff, J.; Dagdigian, P.; Wharton, L. Electric resonance spectrum of NaLi. *J. Chem. Phys.* **1972**, *57*, 710–714.

(211) Mardirossian, N.; Head-Gordon, M. Characterizing and understanding the remarkably slow basis set convergence of several Minnesota density functionals for intermolecular interaction energies. *J. Chem. Theory Comput.* **2013**, *9*, 4453–4461.

(212) Coulson, C. A.; Fischer, I. XXXIV. Notes on the molecular orbital treatment of the hydrogen molecule. *Philos. Mag.* **1949**, *40*, 386–393.

(213) Gwaltney, S. R.; Sherrill, C. D.; Head-Gordon, M.; Krylov, A. I. Second-order perturbation corrections to singles and doubles coupled-cluster methods: General theory and application to the valence optimized doubles model. *J. Chem. Phys.* **2000**, *113*, 3548–3560.

(214) Thakkar, A. J.; Wu, T. How well do static electronic dipole polarizabilities from gas-phase experiments compare with density functional and MP2 computations? *J. Chem. Phys.* **2015**, *143*, 144302.

(215) Rezac, J.; Hobza, P. Describing noncovalent interactions beyond the common approximations: How accurate is the "gold standard," CCSD (T) at the complete basis set limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.

(216) Lonsdale, D. R.; Goerigk, L. The one-electron self-interaction error in 74 density functional approximations: a case study on hydrogenic mono-and dinuclear systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 15805–15830.

(217) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into current limitations of density functional theory. *Science* **2008**, *321*, 792–794.

(218) Hait, D.; Rettig, A.; Head-Gordon, M. Well-behaved versus ill-behaved density functionals for single bond dissociation: Separating success from disaster functional by functional for stretched H2. *J. Chem. Phys.* **2019**, *150*, 094115.

(219) Dohm, S.; Hansen, A.; Steinmetz, M.; Grimme; Stefan; Checinski, M. P. Comprehensive Thermochemical Benchmark Set of Realistic Closed-Shell Metal Organic Reactions. *J. Chem. Theor. Comput.* **2018**, *14*, 2596–2608.

(220) Chan, B.; Gill, P. M. W.; Kimura, M. Assessment of DFT Methods for Transition Metals with the TMC151 Compilation of Data Sets and Comparison with Accuracies for Main-Group Chemistry. *J. Chem. Theor. Comput.* **2019**, *15*, 3610–3622.

(221) Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller–Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 7473–7489.

(222) Leininger, M. L.; Allen, W. D.; Schaefer III, H. F.; Sherrill, C. D. Is Møller–Plesset perturbation theory a convergent ab initio method? *J. Chem. Phys.* **2000**, *112*, 9213–9222.

(223) Hait, D.; Tubman, N. M.; Levine, D. S.; Whaley, K. B.; Head-Gordon, M. What Levels of Coupled Cluster Theory Are Appropriate for Transition Metal Systems? A Study Using Near-Exact Quantum Chemical Values for 3d Transition Metal Binary Compounds. *J. Chem. Theory Comput.* **2019**, *15*, 5370–5385.

(224) Johnson, A. I. et al. The effect of self-interaction error on electrostatic dipoles calculated using density functional theory. *J. Chem. Phys.* **2019**, *151*, 174106.

(225) Grotjahn, R.; Lauter, G. J.; Haasler, M.; Kaupp, M. Evaluation of Local Hybrid Functionals for Electric Properties: Dipole Moments and Static and Dynamic Polarizabilities. *J. Phys. Chem A* **2020**, *124*, 8346–8358.

(226) Hait, D.; Head-Gordon, M. How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry. *Phys. Chem. Chem. Phys.* **2018**, *20*, 19800–19810.

(227) Withanage, K. P. et al. Self-interaction-free electric dipole polarizabilities for atoms and their ions using the Fermi-Löwdin self-interaction correction. *Phys. Rev. A* **2019**, *100*, 012505.

(228) Hait, D.; Rettig, A.; Head-Gordon, M. Beyond the Coulson-Fischer point: Characterizing single excitation CI and TDDFT for excited states in single bond dissociations. *Phys. Chem. Chem. Phys.* **2019**, *21*, 21761–21775.

(229) Cohen, A.; Tantirungrotechai, Y. Molecular electric properties: an assessment of recently developed functionals. *Chem. Phys. Lett.* **1999**, *299*, 465–472.

(230) De Proft, F.; Tielens, F.; Geerlings, P. Performance and basis set dependence of density functional theory dipole and quadrupole moments. *J. Mol. Struct. Theochem* **2000**, *506*, 1–8.

(231) Hofinger, S.; Wendland, M. Method/basis set dependence of the traceless quadrupole moment calculation for N-2, CO2, SO2, HCl, CO, NH3, PH3, HF, and H2O. *Int. J. Quantum Chem.* **2002**, *86*, 199–217.

(232) Matta, C. F. How Dependent are Molecular and Atomic Properties on the Electronic Structure Method? Comparison of Hartree-Fock, DFT, and MP2 on a Biologically Relevant Set of Molecules. *J. Comput. Chem.* **2010**, *31*, 1297–1311.

(233) Halkier, A.; Koch, H.; Christiansen, O.; Jorgensen, P.; Helgaker, T. First-order one-electron properties in the integral-direct coupled cluster singles and doubles model. *J. Chem. Phys.* **1997**, *107*, 849–866.

(234) Zhang, J.; Calvin, J. A.; Valeev, E. F. Anatomy of molecular properties evaluated with explicitly correlated electronic wave functions. *Mol. Phys.* **2016**, *114*, 2894–2909.

(235) Bokhan, D.; Trubnikov, D. N.; Bartlett, R. J. Electric multipole moments calculation with explicitly correlated coupled-cluster wavefunctions. *J. Chem. Phys.* **2016**, *144*, 234107.

(236) Hollett, J. W.; Kelly, A.; Poirier, R. A. Quantum mechanical size and steric hindrance. *J. Phys. Chem. A* **2006**, *110*, 13884–13888.

(237) Lee, J.; Bertels, L. W.; Small, D. W.; Head-Gordon, M. Kohn-sham density functional theory with complex, spin-restricted orbitals: Accessing a new class of densities without the symmetry dilemma. *Phys. Rev. Lett.* **2019**, *123*, 113001.

(238) Brakestad, A. et al. Static polarizabilities at the basis set limit: A benchmark of 124 species. *J. Chem. Theory Comput.* **2020**, *16*, 4874–4882.

(239) Jensen, F. Estimating the Hartree—Fock limit from finite basis set calculations. *Theor. Chem. Acc.* **2005**, *113*, 267–273.

(240) Karton, A.; Martin, J. M. Comment on:"Estimating the Hartree–Fock limit from finite basis set calculations"[Jensen F (2005) Theor Chem Acc 113: 267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.

(241) Jensen, F. Describing anions by density functional theory: fractional electron affinity. *J. Chem. Theory Comput.* **2010**, *6*, 2726–2735.

(242) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The m PW and m PW1PW models. *J. Chem. Phys.* **1998**, *108*, 664–675.

(243) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112.

(244) Tubman, N. M.; Freeman, C. D.; Levine, D. S.; Hait, D.; Head-Gordon, M.; Whaley, K. B. Modern approaches to exact diagonalization and selected configuration interaction with the adaptive sampling CI method. *J. Chem. Theory. Comput.* **2020**, *16*, 2139–2159.

(245) Gaiduk, A. P.; Mizzi, D.; Staroverov, V. N. Self-interaction correction scheme for approximate Kohn-Sham potentials. *Phys. Rev. A* **2012**, *86*, 052518.

(246) Gritsenko, O.; Mentel, Ł.; Baerends, E. On the errors of local density (LDA) and generalized gradient (GGA) approximations to the Kohn-Sham potential and orbital energies. *J. Chem. Phys.* **2016**, *144*, 204114.

(247) Head-Gordon, M. Characterizing unpaired electrons from the one-particle density matrix. *Chem. Phys. Lett.* **2003**, *372*, 508–511.

(248) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3094.

(249) Jensen, F., *Introduction to computational chemistry*; John wiley & sons: 2017.

(250) Runge, E.; Gross, E. K. U. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **1984**, *52*, 997.

(251) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009–4037.

(252) Maekawa, S.; Moorthi, K. Polarizabilities from long-range corrected DFT calculations. *J. Chem. Eng. Data* **2014**, *59*, 3160–3166.

(253) Hohm, U. Experimental static dipole–dipole polarizabilities of molecules. *J. Mol. Struct.* **2013**, *1054*, 282–292.

(254) Sałek, P.; Helgaker, T.; Vahtras, O.; Ågren, H.; Jonsson, D.; Gauss, J. A comparison of density-functional-theory and coupled-cluster frequency-dependent polarizabilities and hyperpolarizabilities. *Mol. Phys.* **2005**, *103*, 439–450.

(255) Hammond, J. R.; Govind, N.; Kowalski, K.; Autschbach, J.; Xantheas, S. S. Accurate dipole polarizabilities for water clusters n= 2–12 at the coupled-cluster level of theory and benchmarking of various density functionals. *J. Chem. Phys.* **2009**, *131*, 214103.

(256) Karne, A. S.; Vaval, N.; Pal, S.; Vásquez-Pérez, J. M.; Köster, A. M.; Calaminici, P. Systematic comparison of DFT and CCSD dipole moments, polarizabilities and hyperpolarizabilities. *Chem. Phys. Lett.* **2015**, *635*, 168–173.

(257) Hu, H.-S. et al. Theoretical studies of the global minima and polarizabilities of small lithium clusters. *Chem. Phys. Lett.* **2016**, *644*, 235–242.

(258) Wu, T.; Kalugina, Y. N.; Thakkar, A. J. Choosing a density functional for static molecular polarizabilities. *Chem. Phys. Lett.* **2015**, *635*, 257–261.

(259) Su, N. Q.; Zhu, Z.; Xu, X. Doubly hybrid density functionals that correctly describe both density and energy for atoms. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, 201713047.

(260) Champagne, B. et al. Assessment of conventional density functional schemes for computing the polarizabilities and hyperpolarizabilities of conjugated oligomers: An ab initio investigation of polyacetylene chains. *J. Chem. Phys.* **1998**, *109*, 10489–10498.

(261) Kirtman, B.; Lacivita, V.; Dovesi, R.; Reis, H. Electric field polarization in conventional density functional theory: From quasilinear to two-dimensional and three-dimensional extended systems. *J. Chem. Phys.* **2011**, *135*, 154101.

(262) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The m PW and m PW1PW models. *J. Chem. Phys.* **1998**, *108*, 664–675.

(263) Bernard, Y. A.; Shao, Y.; Krylov, A. I. General formulation of spin-flip time-dependent density functional theory using non-collinear kernels: Theory, implementation, and benchmarks. *J. Chem. Phys.* **2012**, *136*, 204103.

(264) Ziegler, T.; Rauk, A.; Baerends, E. J. On the calculation of multiplet energies by the Hartree-Fock-Slater method. *Theoretica chimica acta* **1977**, *43*, 261–271.

(265) Hait, D.; Zhu, T.; McMahon, D. P.; Van Voorhis, T. Prediction of Excited-State Energies and Singlet–Triplet Gaps of Charge-Transfer States Using a Restricted Open-Shell Kohn–Sham Approach. *J. Chem. Theory Comput.* **2016**, *12*, 3353–3359.

(266) Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

(267) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

(268) Oviedo, M. B.; Ilawe, N. V.; Wong, B. M. Polarizabilities of π-Conjugated Chains Revisited: Improved Results from Broken-Symmetry Range-Separated DFT and New CCSD (T) Benchmarks. *J. Chem. Theory Comput.* **2016**, *12*, 3593–3602.

(269) Peverati, R.; Head-Gordon, M. Orbital optimized double-hybrid density functionals. *J. Chem. Phys.* **2013**, *139*, 024110.

(270) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(271) Johnson, E. R. A density functional for strong correlation in atoms. *J. Chem. Phys.* **2013**, *139*, 074110.

(272) Kong, J.; Proynov, E. Density functional model for nondynamic and strong correlation. *J. Chem. Theory Comput.* **2015**, *12*, 133–143.

(273) Laqua, H.; Kussmann, J.; Ochsenfeld, C. Communication: Density functional theory model for multi-reference systems based on the exact-exchange hole normalization. *J. Chem. Phys.* **2018**, *148*, 121101.

(274) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I. Binding energy curves from nonempirical density functionals. I. Covalent bonds in closed-shell and radical molecules. *J. Phys. Chem. A* **2005**, *109*, 11006–11014.

(275) Stück, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *J. Chem. Phys.* **2013**, *139*, 244109.

(276) Lee, J.; Head-Gordon, M. Regularized Orbital-Optimized Second-Order Møller–Plesset Perturbation Theory: A Reliable Fifth-Order-Scaling Electron Correlation Model with Orbital Energy Dependent Regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

(277) Thouless, D. Stability conditions and nuclear rotations in the Hartree-Fock theory. *Nuclear Physics* **1960**, *21*, 225–232.

(278) Kim, M.-C.; Sim, E.; Burke, K. Understanding and reducing errors in density functional calculations. *Phys. Rev. Lett.* **2013**, *111*, 073003.

(279) Kim, M.-C.; Sim, E.; Burke, K. Ions in solution: Density corrected density functional theory (DC-DFT). *J. Chem. Phys.* **2014**, *140*, 18A528.

(280) Van Voorhis, T.; Head-Gordon, M. Benchmark variational coupled cluster doubles results. *J. Chem. Phys.* **2000**, *113*, 8873–8879.

(281) Fuchs, M.; Niquet, Y.-M.; Gonze, X.; Burke, K. Describing static correlation in bond dissociation by Kohn–Sham density functional theory. *J. Chem. Phys.* **2005**, *122*, 094116.

(282) Peach, M. J. G.; Teale, A. M.; Tozer, D. J. Modeling the adiabatic connection in H2. *J. Chem. Phys.* **2007**, *126*, 244104.

(283) Dwyer, A. D.; Tozer, D. J. Dispersion, static correlation, and delocalisation errors in density functional theory: An electrostatic theorem perspective. *J. Chem. Phys.* **2011**, *135*, 164110.

(284) Bauernschmitt, R.; Ahlrichs, R. Stability analysis for solutions of the closed shell Kohn–Sham equation. *J. Chem. Phys.* **1996**, *104*, 9047–9052.

(285) Per, M. C.; Fletcher, E. K.; Cleland, D. M. Density functional orbitals in quantum Monte Carlo: The importance of accurate densities. *J. Chem. Phys.* **2019**, *150*, 184101.

(286) Schauperl, M. et al. Non-bonded force field model with advanced restrained electrostatic potential charges (RESP2). *Commun. Chem.* **2020**, *3*, 1–11.

(287) Fedyshyn, O. et al. Spectroscopic and computational study of a new thiazolylazonaphthol dye 1-[(5-(3-nitrobenzyl)-1, 3-thiazol-2-yl) diazenyl] naphthalen-2-ol. *J. Mol. Liq.* **2020**, *304*, 112713.

(288) Raiteri, P.; Kraus, P.; Gale, J. D. Molecular dynamics simulations of liquid–liquid interfaces in an electric field: The water–1, 2-dichloroethane interface. *J. Chem. Phys.* **2020**, *153*, 164714.

(289) Matlock, M. K. et al. Deep Learning Coordinate-Free Quantum Chemistry. *J. Phys. Chem A* **2021**, *125*, 8978–8986.

(290) Jorge, M.; Milne, A. W.; Barrera, M. C.; Gomes, J. R. New force-field for organosilicon molecules in the liquid phase. *ACS Phys. Chem. Au* **2021**, *1*, 54–69.

(291) Zhou, A.; Schauperl, M.; Nerenberg, P. S. Benchmarking electronic structure methods for accurate fixed-charge electrostatic models. *J. Chem. Inf. Model.* **2019**, *60*, 249–258.

(292) Yang, X.; Liu, C.; Walker, B. D.; Ren, P. Accurate description of molecular dipole surface with charge flux implemented for molecular mechanics. *J. Chem. Phys.* **2020**, *153*, 064103.

(293) Elliott, J. D.; Troisi, A.; Carbone, P. A qm/md coupling method to model the ion-induced polarization of graphene. *J. Chem. Theory Comput.* **2020**, *16*, 5253–5263.

(294) Carmona-Espındola, J.; Gázquez, J. L.; Vela, A.; Trickey, S. Generalized gradient approximation exchange energy functional with near-best semilocal performance. *J. Chem. Theory Comput.* **2018**, *15*, 303–310.

(295) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. Regularized second-order Møller–Plesset theory: A more accurate alternative to conventional MP2 for noncovalent interactions and transition metal thermochemistry for the same computational cost. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.

(296) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(297) Chrayteh, A.; Blondel, A.; Loos, P.-F.; Jacquemin, D. Mountaineering Strategy to Excited States: Highly Accurate Oscillator Strengths and Dipole Moments of Small Molecules. *J. Chem. Theory Comput.* **2020**, *17*, 416–438.

(298) Sarkar, R.; Boggio-Pasqua, M.; Loos, P.-F.; Jacquemin, D. Benchmarking TD-DFT and wave function methods for oscillator strengths and excited-state dipole moments. *J. Chem. Theory Comput.* **2021**, *17*, 1117–1132.

(299) Lykhin, A. O.; Truhlar, D. G.; Gagliardi, L. Dipole Moment Calculations Using Multiconfiguration Pair-Density Functional Theory and Hybrid Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput* **2021**, *17*, 7586–7601.

(300) Wilkins, D. M.; Grisafi, A.; Yang, Y.; Lao, K. U.; DiStasio, R. A.; Ceriotti, M. Accurate molecular polarizabilities with coupled cluster theory and machine learning. *Proc. Nat. Acad. Sci.* **2019**, *116*, 3401–3406.

(301) Turro, N. J.; Ramamurthy, V.; Ramamurthy, V.; Scaiano, J. C., *Principles of molecular photochemistry: an introduction*; University science books: 2009.

(302) Loos, P.-F.; Scemama, A.; Jacquemin, D. The quest for highly accurate excitation energies: A computational perspective. *J. Phys. Chem. Lett.* **2020**, *11*, 2374–2383.

(303) Singh, R.; Deb, B. M. Developments in excited-state density functional theory. *Phys. Rep.* **1999**, *311*, 47–94.

(304) Ghosh, S.; Verma, P.; Cramer, C. J.; Gagliardi, L.; Truhlar, D. G. Combining wave function methods with density functional theory for excited states. *Chem. Rev.* **2018**, *118*, 7249–7292.

(305) Ferré, N.; Filatov, M.; Huix-Rotllant, M.; Adamo, C., *Density-functional methods for excited states*; Springer: 2016.

(306) Casida, M. E. In *Recent Advances In Density Functional Methods: (Part I)*; World Scientific: 1995, pp 155–192.

(307) Seidl, A.; Görling, A.; Vogl, P.; Majewski, J.; Levy, M. Generalized Kohn-Sham schemes and the band-gap problem. *Phys. Rev. B* **1996**, *53*, 3764.

(308) Heinrichs, J. New derivation of time-dependent Hartree-Fock theory. *Chemical Physics Lett.* **1968**, *2*, 315–318.

(309) Blase, X.; Duchemin, I.; Jacquemin, D.; Loos, P.-F. The Bethe–Salpeter equation formalism: From physics to chemistry. *J. Phys. Chem. Lett.* **2020**, *11*, 7371–7382.

(310) Van Setten, M. J. et al. GW 100: Benchmarking G 0 W 0 for molecular systems. *J. Chem. Theory Comput.* **2015**, *11*, 5665–5687.

(311) Rangel, T.; Hamed, S. M.; Bruneval, F.; Neaton, J. B. Evaluating the GW approximation with CCSD (T) for charged excitations across the oligoacenes. *J. Chem. Theory Comput.* **2016**, *12*, 2834–2842.

(312) Véril, M.; Romaniello, P.; Berger, J.; Loos, P.-F. Unphysical discontinuities in GW methods. *J. Chem. Theory Comput.* **2018**, *14*, 5220–5228.

(313) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

(314) Tamm, I. In *Selected Papers*; Springer: 1991, pp 157–174.

(315) Dancoff, S. Non-adiabatic meson theory of nuclear forces. *Phys. Rev.* **1950**, *78*, 382.

(316) Cordova, F.; Doriol, L. J.; Ipatov, A.; Casida, M. E.; Filippi, C.; Vela, A. Troubleshooting time-dependent density-functional theory for photochemical applications: Oxirane. *J. Chem. Phys.* **2007**, *127*, 164111.

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

(318) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. *J. Chem. Phys.* **1998**, *109*, 8218–8224.

(319) Liang, J.; Feng, X.; Hait, D.; Head-Gordon, M. Revisiting the performance of time-dependent density functional theory for electronic excitations: Assessment of 43 popular and recently developed functionals from rungs one to four. *J. Chem. Theory. Comput.* **2022**, *Accepted.*

(320) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem* **1992**, *96*, 135–149.

(321) Subotnik, J. Communication: configuration interaction singles has a large systematic bias against charge-transfer states. *J. Chem. Phys.* **2011**, *135*, 071104–071104.

(322) Maurice, D.; Head-Gordon, M. On the nature of electronic transitions in radicals: An extended single excitation configuration interaction method. *J. Phys. Chem.* **1996**, *100*, 6131–6137.

(323) Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **2002**, *117*, 7433–7447.

(324) Jacquemin, D.; Wathelet, V.; Perpete, E. A.; Adamo, C. Extensive TD-DFT benchmark: singlet-excited states of organic molecules. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

(325) Tozer, D. J.; Handy, N. C. On the determination of excitation energies using density functional theory. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2117–2121.

(326) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* **2004**, *120*, 5932–5937.

(327) Levine, B. G.; Ko, C.; Quenneville, J.; MartÍnez, T. J. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* **2006**, *104*, 1039–1051.

(328) Besley, N. A.; Asmuruf, F. A. Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039.

(329) Baruah, T.; Olguin, M.; Zope, R. R. Charge transfer excited state energies by perturbative delta self consistent field method. *J. Chem. Phys.* **2012**, *137*, 084316.

(330) Kronik, L.; Stein, T.; Refaely-Abramson, S.; Baer, R. Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals. *J. Chem. Theory Comput.* **2012**, *8*, 1515–1531.

(331) Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den einzelnen Elektronen eines Atoms. *Physica* **1934**, *1*, 104–113.

(332) Shee, J.; Head-Gordon, M. Predicting Excitation Energies of Twisted Intramolecular Charge-Transfer States with the Time-Dependent Density Functional Theory: Comparison with Experimental Measurements in the Gas Phase and Solvents Ranging from Hexanes to Acetonitrile. *J. Chem. Theory Comput.* **2020**, *16*, 6244–6255.

(333) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals. *J. Phys. Chem A* **2007**, *111*, 8753–8765.

(334) Phillipson, P.; Mulliken, R. Improved Molecular Orbitals (Computations on H2). *J. Chem. Phys.* **1958**, *28*, 1248–1249.

(335) Hunt, W.; Dunning Jr, T.; Goddard III, W. The orthogonality constrained basis set expansion method for treating off-diagonal lagrange multipliers in calculations of electronic wave functions. *Chem. Phys. Lett.* **1969**, *3*, 606–610.

(336) Basch, H.; McKoy, V. Interpretation of Open-Shell SCF Calculations on the T and V States of Ethylene. *J. Chem. Phys.* **1970**, *53*, 1628–1637.

(337) Bagus, P.; Bennett, B. Singlet–triplet splittings as obtained from the Xα-scattered wave method: A theoretical analysis. *Int. J. Quantum Chem* **1975**, *9*, 143–148.

(338) Cheng, C.-L.; Wu, Q.; Van Voorhis, T. Rydberg energies using excited state density functional theory. *J. Chem. Phys.* **2008**, *129*, 124112.

(339) Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained density functional theory. *Chem. Rev.* **2012**, *112*, 321–370.

(340) Bagus, P. S. Self-consistent-field wave functions for hole states of some Ne-like and Ar-like ions. *Phys. Rev.* **1965**, *139*, A619.

(341) Gilbert, A. T.; Besley, N. A.; Gill, P. M. Self-consistent field calculations of excited states using the maximum overlap method (MOM). *J. Phys. Chem A* **2008**, *112*, 13164–13171.

(342) Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.

(343) Mewes, J.-M.; Jovanović, V.; Marian, C. M.; Dreuw, A. On the molecular mechanism of non-radiative decay of nitrobenzene and the unforeseen challenges this simple molecule holds for electronic structure theory. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12393–12406.

(344) Barca, G. M.; Gilbert, A. T.; Gill, P. M. Simple Models for Difficult Electronic Excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501–1509.

(345) Hait, D.; Head-Gordon, M. Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory. *J. Chem. Theory Comput.* **2020**, *16*, 1699–1710.

(346) Saunders, V.; Hillier, I. A "Level–Shifting" method for converging closed shell Hartree–Fock wave functions. *Int. J. Quantum Chem* **1973**, *7*, 699–705.

(347) Bacskay, G. B. A quadratically convergent Hartree—Fock (QC-SCF) method. Application to closed shell systems. *Chem. Phys.* **1981**, *61*, 385–404.

(348) Van Voorhis, T.; Head-Gordon, M. A geometric approach to direct minimization. *Mol. Phys.* **2002**, *100*, 1713–1721.

(349) Weinstein, D. Modified Ritz method. *Proc. Natl. Acad. Sci. U.S.A.* **1934**, *20*, 529.

(350) Umrigar, C.; Wilson, K.; Wilkins, J. Optimized trial wave functions for quantum Monte Carlo calculations. *Phys. Rev. Lett.* **1988**, *60*, 1719.

(351) Zhao, L.; Neuscamman, E. An efficient variational principle for the direct optimization of excited states. *J. Chem. Theory Comput.* **2016**, *12*, 3436–3440.

(352) Shea, J. A.; Neuscamman, E. Communication: A mean field platform for excited state quantum chemistry. *J. Chem. Phys.* **2018**, *149*, 081101.

(353) Ye, H.-Z.; Welborn, M.; Ricke, N. D.; Van Voorhis, T. σ-SCF: A direct energy-targeting method to mean-field excited states. *J. Chem. Phys.* **2017**, *147*, 214104.

(354) Ye, H.-Z.; Van, T. V. Half-Projected σ Self-Consistent Field For Electronic Excited States. *J. Chem. Theory Comput.* **2019**, *15*, 2954–2965.

(355) Loos, P.-F.; Boggio-Pasqua, M.; Scemama, A.; Caffarel, M.; Jacquemin, D. Reference energies for double excitations. *J. Chem. Theory Comput.* **2019**, *15*, 1939–1956.

(356) Hait, D.; Haugen, E. A.; Yang, Z.; Oosterbaan, K. J.; Leone, S. R.; Head-Gordon, M. Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations. *J. Chem. Phys.* **2020**, *153*, 134108.

(357) Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. A spin correction procedure for unrestricted Hartree-Fock and Möller-Plesset wavefunctions for singlet diradicals and polyradicals. *Chem. Phys. Lett.* **1988**, *149*, 537–542.

(358) Kowalczyk, T.; Yost, S. R.; Voorhis, T. V. Assessment of the ΔSCF density functional theory approach for electronic excitations in organic dyes. *J. Chem. Phys.* **2011**, *134*, 054128.

(359) Han, W.-G.; Lovell, T.; Liu, T.; Noodleman, L. Density Functional Studies of the Ground-and Excited-State Potential-Energy Curves of Stilbene cis–trans Isomerization. *ChemPhysChem* **2002**, *3*, 167–178.

(360) Liu, T. et al. Density functional vertical self-consistent reaction field theory for solvatochromism studies of solvent-sensitive dyes. *J. Phys. Chem. A* **2004**, *108*, 3545–3555.

(361) Frank, I.; Hutter, J.; Marx, D.; Parrinello, M. Molecular dynamics in low-spin excited states. *J. Chem. Phys.* **1998**, *108*, 4060–4069.

(362) Filatov, M.; Shaik, S. A spin-restricted ensemble-referenced Kohn–Sham method and its application to diradicaloid situations. *Chem. Phys. Lett.* **1999**, *304*, 429–437.

(363) Kowalczyk, T.; Tsuchimochi, T.; Chen, P.-T.; Top, L.; Van Voorhis, T. Excitation energies and Stokes shifts from a restricted open-shell Kohn-Sham approach. *J. Chem. Phys.* **2013**, *138*, 164101.

(364) Russo, T. V.; Martin, R. L.; Hay, P. J. Density functional calculations on first-row transition metals. *J. Chem. Phys.* **1994**, *101*, 7729–7737.

(365) Filatov, M.; Shaik, S. Spin-restricted density functional approach to the open-shell problem. *Chem. Phys. Lett.* **1998**, *288*, 689–697.

(366) Hait, D.; Head-Gordon, M. Orbital optimized density functional theory for electronic excited states. *J. Chem. Phys. Lett.* **2021**, *12*, 4517–4529.

(367) Gaudoin, R.; Burke, K. Lack of Hohenberg-Kohn theorem for excited states. *Phys. Rev. Lett.* **2004**, *93*, 173001.

(368) Ayers, P. W.; Levy, M.; Nagy, A. Time-independent density-functional theory for excited states of Coulomb systems. *Phys. Rev. A* **2012**, *85*, 042518.

(369) Shea, J. A.; Gwin, E.; Neuscamman, E. A generalized variational principle with applications to excited state mean field theory. *J. Chem. Theory Comput.* **2020**, *16*, 1526–1540.

(370) Wilson, S. A. et al. X-ray absorption spectroscopic investigation of the electronic structure differences in solution and crystalline oxyhemoglobin. *Proceedings of the National Academy of Sciences* **2013**, *110*, 16333–16338.

(371) Hähner, G.; Hofer, R.; Klingenfuss, I. Order and orientation in self-assembled long chain alkanephosphate monolayers adsorbed on metal oxide surfaces. *Langmuir* **2001**, *17*, 7047–7052.

(372) Guo, Q. et al. Electronic structure of GaInN semiconductors investigated by x-ray absorption spectroscopy. *Applied Physics Lett.* **2011**, *98*, 181901.

(373) Chergui, M.; Collet, E. Photoinduced structural dynamics of molecular systems mapped by time-resolved X-ray methods. *Chem. Rev.* **2017**, *117*, 11025–11065.

(374) Bhattacherjee, A.; Leone, S. R. Ultrafast X-ray Transient Absorption Spectroscopy of Gas-Phase Photochemical Reactions: A New Universal Probe of Photoinduced Molecular Dynamics. *Acc. Chem. Res.* **2018**, *51*, 3203–3211.

(375) Schnorr, K. et al. Tracing the 267 nm-induced radical formation in dimethyl disulfide using time-resolved X-ray absorption spectroscopy. *J. Phys. Chem Lett.* **2019**, *10*, 1382–1387.

(376) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for the calculation of core-excited states. *J. Chem. Phys.* **2018**, *149*, 044116.

(377) Besley, N. A.; Peach, M. J.; Tozer, D. J. Time-dependent density functional theory calculations of near-edge X-ray absorption fine structure with short-range corrected functionals. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10350–10358.

(378) Frati, F. et al. Coupled cluster study of the x-ray absorption spectra of formaldehyde derivatives at the oxygen, carbon, and fluorine K-edges. *J. Chem. Phys.* **2019**, *151*, 064107.

(379) Marques, M. A.; Gross, E. K. Time-dependent density functional theory. *Annu. Rev. Phys. Chem.* **2004**, *55*, 427–455.

(380) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martınez, T. J. Excited-state electronic structure with configuration interaction singles and Tamm–Dancoff time-dependent density functional theory on graphical processing units. *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.

(381) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.

(382) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.

(383) Roos, B. O.; Taylor, P. R.; Si, P. E., et al. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157–173.

(384) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem* **1990**, *94*, 5483–5488.

(385) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

(386) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. A long-range-corrected time-dependent density functional theory. *J. Chem. Phys.* **2004**, *120*, 8425–8433.

(387) Peach, M. J.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.* **2008**, *128*, 044118.

(388) Sun, H.; Zhong, C.; Bredas, J.-L. Reliable prediction with tuned range-separated functionals of the singlet–triplet gap in organic emitters for thermally activated delayed fluorescence. *J. Chem. Theory Comput.* **2015**, *11*, 3851–3858.

(389) Bauernschmitt, R.; Ahlrichs, R. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **1996**, *256*, 454–464.

(390) Peach, M. J.; Williamson, M. J.; Tozer, D. J. Influence of triplet instabilities in TDDFT. *J. Chem. Theory Comput.* **2011**, *7*, 3578–3585.

(391) Čižek, J.; Paldus, J. Stability Conditions for the Solutions of the Hartree—Fock Equations for Atomic and Molecular Systems. Application to the Pi-Electron Model of Cyclic Polyenes. *J. Chem. Phys.* **1967**, *47*, 3976–3985.

(392) Peach, M. J.; Tozer, D. J. Overcoming low orbital overlap and triplet instability problems in TDDFT. *J. Phys. Chem A* **2012**, *116*, 9783–9789.

(393) Giesbertz, K.; Baerends, E. Failure of time-dependent density functional theory for excited state surfaces in case of homolytic bond dissociation. *Chem. Phys. Lett.* **2008**, *461*, 338–342.

(394) Thom, A. J.; Head-Gordon, M. Hartree–Fock solutions as a quasidiabatic basis for nonorthogonal configuration interaction. *J. Chem. Phys.* **2009**, *131*, 124113.

(395) Sundstrom, E. J.; Head-Gordon, M. Non-orthogonal configuration interaction for the calculation of multielectron excited states. *J. Chem. Phys.* **2014**, *140*, 114103.

(396) Wu, Q.; Cheng, C.-L.; Van Voorhis, T. Configuration interaction based on constrained density functional theory: A multireference method. *J. Chem. Phys.* **2007**, *127*, 164119.

(397) Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained density functional theory. *J. Mol. Spec. reviews* **2011**, *112*, 321–370.

(398) Myneni, H.; Casida, M. E. On the calculation of $\Delta{<}S^2{>}$ for electronic excitations in time-dependent density-functional theory. *Comput. Phys. Comm.* **2017**, *213*, 72–91.

(399) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. A doubles correction to electronic excited states from configuration interaction in the space of single substitutions. *Chem. Phys. Lett.* **1994**, *219*, 21–29.

(400) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

(401) Hehre, W. J.; Stewart, R. F.; Pople, J. A. self-consistent molecular-orbital methods. i. use of gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.* **1969**, *51*, 2657–2664.

(402) Cui, Y.; Bulik, I. W.; Jiménez-Hoyos, C. A.; Henderson, T. M.; Scuseria, G. E. Proper and improper zero energy modes in Hartree-Fock theory and their relevance for symmetry breaking and restoration. *J. Chem. Phys.* **2013**, *139*, 154107.

(403) Burton, H. G.; Thom, A. J. Holomorphic Hartree–Fock theory: an inherently multireference approach. *J. Chem. Theory Comput.* **2015**, *12*, 167–173.

(404) Elliott, P.; Goldson, S.; Canahui, C.; Maitra, N. T. Perspectives on double-excitations in TDDFT. *Chem. Phys.* **2011**, *391*, 110–119.

(405) Dreuw, A.; Head-Gordon, M. Failure of time-dependent density functional theory for long-range charge-transfer excited states: the zincbacteriochlorin- bacteriochlorin and bacteriochlorophyll- spheroidene complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

(406) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold. *J. Chem. Phys.* **1998**, *108*, 4439–4449.

(407) Ziegler, T.; Seth, M.; Krykunov, M.; Autschbach, J. A revised electronic Hessian for approximate time-dependent density functional theory. *J. Chem. Phys.* **2008**, *129*, 184114.

(408) Zhekova, H. R.; Seth, M.; Ziegler, T. A perspective on the relative merits of time-dependent and time-independent density functional theory in studies of the electron spectra due to transition metal complexes. An illustration through applications to copper tetrachloride and plastocyanin. *Int. J. Quantum Chem* **2014**, *114*, 1019–1029.

(409) Thom, A. J.; Head-Gordon, M. Locating multiple self-consistent field solutions: an approach inspired by metadynamics. *Phys. Rev. Lett.* **2008**, *101*, 193001.

(410) Liu, X.; Fatehi, S.; Shao, Y.; Veldkamp, B.; Subotnik, J. Communication: Adjusting charge transfer state energies for configuration interaction singles: without any parameterization and with minimal cost. *J. Chem. Phys.* **2012**, *136*, 161101–161101.

(411) Krykunov, M.; Seth, M.; Ziegler, T. Introducing constricted variational density functional theory in its relaxed self-consistent formulation (RSCF-CV-DFT) as an alternative to adiabatic time dependent density functional theory for studies of charge transfer transitions. *J. Chem. Phys.* **2014**, *140*, 18A502.

(412) Tran, L. N.; Shea, J. A. R.; Neuscamman, E. Tracking Excited States in Wave Function Optimization Using Density Matrices and Variational Principles. *J. Chem. Theory Comput.* **2019**, *15*, 4790–4803.

(413) Nite, J.; Jiménez-Hoyos, C. A. Low-Cost Molecular Excited States from a State-Averaged Resonating Hartree–Fock Approach. *J. Chem. Theory Comput.* **2019**, *15*, 5343–5351.

(414) Burton, H. G. A.; Thom, A. J. W. General Approach for Multireference Ground and Excited States Using Nonorthogonal Configuration Interaction. *J. Chem. Theory Comput.* **2019**, *15*, 4851–4861.

(415) Zhao, L.; Neuscamman, E. Density functional extension to excited-state mean-field theory. *J. Chem. Theory Comput.* **2019**, *16*, 164–178.

(416) Besley, N. A.; Gilbert, A. T.; Gill, P. M. Self-consistent-field calculations of core excited states. *J. Chem. Phys.* **2009**, *130*, 124308.

(417) Zhao, Q.; Ioannidis, E. I.; Kulik, H. J. Global and local curvature in density functional theory. *J. Chem. Phys.* **2016**, *145*, 054109.

(418) Bozkaya, U. Orbital-Optimized Second-Order Perturbation Theory with Density-Fitting and Cholesky Decomposition Approximations: An Efficient Implementation. *J. Chem. Theor. Comput.* **2014**, *10*, 2371–2378.

(419) Sherrill, C. D.; Krylov, A. I.; Byrd, E. F.; Head-Gordon, M. Energies and analytic gradients for a coupled-cluster doubles model using variational Brueckner orbitals: application to symmetry breaking in O 4+. *J. Chem. Phys.* **1998**, *109*, 4171–4181.

(420) Sharada, S. M.; Stück, D.; Sundstrom, E. J.; Bell, A. T.; Head-Gordon, M. Wavefunction stability analysis without analytical electronic Hessians: application to orbital-optimised second-order Møller–Plesset theory and VV10-containing density functionals. *Mol. Phys.* **2015**, *113*, 1802–1808.

(421) Deglmann, P.; Furche, F. Efficient characterization of stationary points on potential energy surfaces. *J. Chem. Phys.* **2002**, *117*, 9535–9538.

(422) Reiher, M.; Neugebauer, J. A mode-selective quantum chemical method for tracking molecular vibrations applied to functionalized carbon nanotubes. *J. Chem. Phys.* **2003**, *118*, 1634–1641.

(423) Kaledin, A. L. Gradient-based direct normal-mode analysis. *J. Chem. Phys.* **2005**, *122*, 184106.

(424) Sharada, S. M.; Bell, A. T.; Head-Gordon, M. A finite difference Davidson procedure to sidestep full ab initio hessian calculation: Application to characterization of stationary points and transition state searches. *J. Chem. Phys.* **2014**, *140*, 164115.

(425) Fletcher, R., *Practical methods of optimization*; John Wiley & Sons: 2013.

(426) Broyden, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA J. Appl. Math.* **1970**, *6*, 76–90.

(427) Fletcher, R. A new approach to variable metric algorithms. *The computer journal* **1970**, *13*, 317–322.

(428) Goldfarb, D. A family of variable-metric methods derived by variational means. *Mathematics of computation* **1970**, *24*, 23–26.

(429) Shanno, D. F. Conditioning of quasi-Newton methods for function minimization. *Mathematics of computation* **1970**, *24*, 647–656.

(430) Abadi, M. et al. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp 265–283.

(431) Barca, G. M.; Gilbert, A. T.; Gill, P. M. Excitation number: characterizing multiply excited states. *J. Chem. Theory Comput.* **2017**, *14*, 9–13.

(432) Prlj, A.; Sandoval-Salinas, M. E.; Casanova, D.; Jacquemin, D.; Corminboeuf, C. Low-lying $\pi\pi^*$ states of heteroaromatic molecules: A challenge for excited state methods. *J. Chem. Theory Comput.* **2016**, *12*, 2652–2660.

(433) Cullen, J.; Krykunov, M.; Ziegler, T. The formulation of a self-consistent constricted variational density functional theory for the description of excited states. *Chem. Phys.* **2011**, *391*, 11–18.

(434) Ziegler, T.; Krykunov, M.; Cullen, J. The implementation of a self-consistent constricted variational density functional theory for the description of excited states. *J. Chem. Phys.* **2012**, *136*, 124107.

(435) Evangelista, F. A.; Shushkov, P.; Tully, J. C. Orthogonality constrained density functional theory for electronic excited states. *J. Phys. Chem A* **2013**, *117*, 7378–7392.

(436) Ramos, P.; Pavanello, M. Low-lying excited states by constrained DFT. *J. Chem. Phys.* **2018**, *148*, 144103.

(437) Kendall, R. A.; Dunning Jr, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(438) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(439) Christiansen, O.; Koch, H.; Jo/rgensen, P. Response functions in the CC3 iterative triple excitation model. *J. Chem. Phys.* **1995**, *103*, 7429–7441.

(440) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.* **2018**, *14*, 4360–4379.

(441) Ye, H.-Z.; Van Voorhis, T. Self-consistent Møller-Plesset Perturbation Theory For Excited States. *arXiv preprint arXiv:2008.10777* **2020**.

(442) Armstrong, N. R.; Wang, W.; Alloway, D. M.; Placencia, D.; Ratcliff, E.; Brumbach, M. Organic/organic heterojunctions: organic light emitting diodes and organic photovoltaic devices. *Macromol. Rapid Commun* **2009**, *30*, 717–731.

(443) Kushto, G. P.; Makinen, A. J.; Lane, P. A. Organic photovoltaic cells using group 10 metallophthalocyanine electron donors. *IEEE J. Sel. Top. Quantum Electron* **2010**, *16*, 1552–1559.

(444) Walter, M. G.; Rudine, A. B.; Wamser, C. C. Porphyrins and phthalocyanines in solar photovoltaic cells. *J. Porphyr. Phthalocyanines* **2010**, *14*, 759–792.

(445) Bonnett, R. Photosensitizers of the porphyrin and phthalocyanine series for photodynamic therapy. *Chem. Soc. Rev.* **1995**, *24*, 19–33.

(446) Edwards, L.; Gouterman, M. Porphyrins: XV. Vapor absorption spectra and stability: Phthalocyanines. *J. Mol. Spec.* **1970**, *33*, 292–310.

(447) Nyokong, T.; Gasyna, Z.; Stillman, M. J. Analysis of the absorption and magnetic circular dichroism spectra of zinc phthalocyanine and the. pi.-cation-radical species [ZnPc (1-)]. cntdot.+. *Inorg. Chem.* **1987**, *26*, 1087–1095.

(448) Metcalf, D. H.; VanCott, T. C.; Snyder, S. W.; Schatz, P. N.; Williamson, B. E. Magnetic circularly polarized luminescence of zinc phthalocyanine in an argon matrix. *J. Phys. Chem* **1990**, *94*, 2828–2832.

(449) Mack, J.; Stillman, M. J. Band deconvolution analysis of the absorption and magnetic circular dichroism spectral data of ZnPc (-2) recorded at cryogenic temperatures. *J. Phys. Chem* **1995**, *99*, 7935–7945.

(450) Van Cott, T. C. et al. Magnetic circular dichroism and absorption spectrum of zinc phthalocyanine in an argon matrix between 14700 and 74000 cm-1. *J. Phys. Chem* **1989**, *93*, 2999–3011.

(451) Nguyen, K. A.; Pachter, R. Ground state electronic structures and spectra of zinc complexes of porphyrin, tetraazaporphyrin, tetrabenzoporphyrin, and phthalocyanine: A density functional theory study. *J. Chem. Phys.* **2001**, *114*, 10757–10767.

(452) Ricciardi, G.; Rosa, A.; Baerends, E. Ground and excited states of zinc phthalocyanine studied by density functional methods. *J. Phys. Chem A* **2001**, *105*, 5242–5254.

(453) Peralta, G.; Seth, M.; Zhekova, H.; Ziegler, T. Magnetic circular dichroism of phthalocyanine (M= Mg, Zn) and tetraazaporphyrin (M= Mg, Zn, Ni) metal complexes. A computational study based on time-dependent density functional theory. *Inorg. Chem.* **2008**, *47*, 4185–4198.

(454) Theisen, R. F.; Huang, L.; Fleetham, T.; Adams, J. B.; Li, J. Ground and excited states of zinc phthalocyanine, zinc tetrabenzoporphyrin, and azaporphyrin analogs using DFT and TDDFT with Franck-Condon analysis. *J. Chem. Phys.* **2015**, *142*, 094310.

(455) Wallace, A. J.; Williamson, B. E.; Crittenden, D. L. Coupled cluster calculations provide a one-to-one mapping between calculated and observed transition energies in the electronic absorption spectrum of zinc phthalocyanine. *Int. J. Quantum Chem* **2017**, *117*, e25350.

(456) Zheng, X.; Cheng, L. Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements. *J. Chem. Theory Comput.* **2019**, *15*, 4945–4955.

(457) Lee, J.; Small, D. W.; Head-Gordon, M. Excited states via coupled cluster theory without equation-of-motion methods: Seeking higher roots with application to doubly excited states and double core hole states. *J. Chem. Phys.* **2019**, *151*, 214103.

(458) Yost, S. R.; Head-Gordon, M. Size consistent formulations of the perturb-then-diagonalize Møller-Plesset perturbation theory correction to non-orthogonal configuration interaction. *J. Chem. Phys.* **2016**, *145*, 054105.

(459) Yost, S. R.; Head-Gordon, M. Efficient Implementation of NOCI-MP2 Using the Resolution of the Identity Approximation with Application to Charged Dimers and Long C–C Bonds in Ethane Derivatives. *J. Chem. Theory Comput.* **2018**, *14*, 4791–4805.

(460) Wenzel, J.; Wormit, M.; Dreuw, A. Calculating X-ray absorption spectra of open-shell molecules with the unrestricted algebraic-diagrammatic construction scheme for the polarization propagator. *J. Chem. Theory Comput.* **2014**, *10*, 4583–4598.

(461) Attar, A. R. et al. Femtosecond x-ray spectroscopy of an electrocyclic ring-opening reaction. *Science* **2017**, *356*, 54–59.

(462) Bhattacherjee, A. et al. Photoinduced Heterocyclic Ring Opening of Furfural: Distinct Open-Chain Product Identification by Ultrafast X-ray Transient Absorption Spectroscopy. *J. Am. Chem. Soc.* **2018**, *140*, 12538–12544.

(463) Hanson-Heine, M. W.; George, M. W.; Besley, N. A. A scaled CIS (D) based method for the calculation of valence and core electron ionization energies. *J. Chem. Phys.* **2019**, *151*, 034104.

(464) Seidu, I.; Neville, S. P.; Kleinschmidt, M.; Heil, A.; Marian, C. M.; Schuurman, M. S. The simulation of X-ray absorption spectra from ground and excited electronic states using core-valence separated DFT/MRCI. *J. Chem. Phys.* **2019**, *151*, 144104.

(465) Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. Coupled-cluster response theory for near-edge x-ray-absorption fine structure of atoms and molecules. *Phys. Rev. A* **2012**, *85*, 022507.

(466) Wormit, M. et al. Investigating excited electronic states using the algebraic diagrammatic construction (ADC) approach of the polarisation propagator. *Mol. Phys.* **2014**, *112*, 774–784.

(467) Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **1965**, *139*, A796.

(468) Salpeter, E. E.; Bethe, H. A. A relativistic equation for bound-state problems. *Phys. Rev.* **1951**, *84*, 1232.

(469) Shirley, E. L. Ti 1s pre-edge features in rutile: a Bethe-Salpeter calculation. *J. Electron Spectrosc. Relat. Phenom.* **2004**, *136*, 77–83.

(470) Vinson, J.; Rehr, J.; Kas, J.; Shirley, E. Bethe-Salpeter equation calculations of core excitation spectra. *Phys. Rev. B.* **2011**, *83*, 115106.

(471) Gilmore, K. et al. Efficient implementation of core-excitation Bethe–Salpeter equation calculations. *Comp. Phys. Comm.* **2015**, *197*, 109–117.

(472) Oakley, M. S.; Klobukowski, M. ΔDFT/MIX: A reliable and efficient method for calculating core electron binding energies of large molecules. *J. Electron Spectrosc. Relat. Phenom.* **2018**, *227*, 44–50.

(473) Kahk, J. M.; Lischner, J. Accurate absolute core-electron binding energies of molecules, solids, and surfaces from first-principles calculations. *Phys. Rev. Materials* **2019**, *3*, 100801.

(474) Derricotte, W. D.; Evangelista, F. A. Simulation of X-ray absorption spectra with orthogonality constrained density functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.

(475) Michelitsch, G. S.; Reuter, K. Efficient simulation of near-edge x-ray absorption fine structure (NEXAFS) in density-functional theory: Comparison of core-level constraining approaches. *J. Chem. Phys.* **2019**, *150*, 074104.

(476) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. Direct SCF direct static-exchange calculations of electronic spectra. *Theo. Chem. Acc.* **1997**, *97*, 14–40.

(477) Oosterbaan, K. J.; White, A. F.; Head-Gordon, M. Non-orthogonal configuration interaction with single substitutions for core-excited states: An extension to doublet radicals. *J. Chem. Theory Comput.* **2019**, *15*, 2966–2973.

(478) Leetmaa, M. et al. Are recent water models obtained by fitting diffraction data consistent with infrared/Raman and x-ray absorption spectra? *J. Chem. Phys.* **2006**, *125*, 244510.

(479) Leetmaa, M.; Ljungberg, M.; Lyubartsev, A.; Nilsson, A.; Pettersson, L. G. Theoretical approximations to X-ray absorption spectroscopy of liquid water and ice. *J. Electron Spectrosc. Relat. Phenom.* **2010**, *177*, 135–157.

(480) Hitchcock, A.; Brion, C. Carbon K-shell excitation of C2H2, C2H4, C2H6 and C6H6 by 2.5 keV electron impact. *J. Electron Spectrosc. Relat. Phenom.* **1977**, *10*, 317–330.

(481) Remmers, G. et al. High-resolution K-shell photoabsorption in formaldehyde. *Phys. Rev. A* **1992**, *46*, 3935.

(482) Hitchcock, A.; Brion, C. Inner shell electron energy loss studies of HCN and C2N2. *Chem. Phys.* **1979**, *37*, 319–331.

(483) Prince, K. C.; Richter, R.; de Simone, M.; Alagia, M.; Coreno, M. Near edge X-ray absorption spectra of some small polyatomic molecules. *J. Phys. Chem A* **2003**, *107*, 1955–1963.

(484) Domke, M. et al. Carbon and oxygen K-edge photoionization of the CO molecule. *Chem. Phys. Lett.* **1990**, *173*, 122–128.

(485) Schirmer, J. et al. K-shell excitation of the water, ammonia, and methane molecules using high-resolution photoabsorption spectroscopy. *Phys. Rev. A* **1993**, *47*, 1136.

(486) Robin, M. B.; Ishii, I.; McLaren, R.; Hitchcock, A. P. Fluorination effects on the inner-shell spectra of unsaturated molecules. *J. Electron Spectrosc. Relat. Phenom.* **1988**, *47*, 53–92.

(487) Prince, K.; Avaldi, L.; Coreno, M.; Camilloni, R.; De Simone, M. Vibrational structure of core to Rydberg state excitations of carbon dioxide and dinitrogen oxide. *J. Phys. B: At. Mol. Opt. Phys.* **1999**, *32*, 2551.

(488) Apen, E.; Hitchcock, A. P.; Gland, J. L. Experimental studies of the core excitation of imidazole, 4, 5-dicyanoimidazole, and s-triazine. *J. Phys. Chem* **1993**, *97*, 6859–6866.

(489) Myhre, R. H. et al. A theoretical and experimental benchmark study of core-excited states in nitrogen. *J. Chem. Phys.* **2018**, *148*, 064106.

(490) Plekan, O. et al. An X-ray absorption study of glycine, methionine and proline. *J. Electron Spectrosc. Relat. Phenom.* **2007**, *155*, 47–53.

(491) Pavlychev, A.; Hallmeier, K.; Hennig, C.; Hennig, L.; Szargan, R. Nitrogen K-shell excitations in complex molecules and polypyrrole. *Chem. Phys.* **1995**, *201*, 547–555.

(492) Duflot, D.; Flament, J.-P.; Giuliani, A.; Heinesch, J.; Hubin-Franskin, M.-J. Core shell excitation of furan at the O 1s and C 1s edges: An experimental and ab initio study. *J. Chem. Phys.* **2003**, *119*, 8946–8955.

(493) Hitchcock, A.; Brion, C. K-shell excitation of HF and F2 studied by electron energy-loss spectroscopy. *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 4399.

(494) Jolly, W.; Bomben, K.; Eyermann, C. Core-electron binding energies for gaseous atoms and molecules. *Atomic Data and Nuclear Data Tables* **1984**, *31*, 433–493.

(495) Takahashi, O. Relativistic corrections for single-and double-core excitation at the K- and L-edges from Li to Kr. *Computational and Theoretical Chemistry* **2017**, *1102*, 80–86.

(496) Oosterbaan, K. J.; White, A. F.; Hait, D.; Head-Gordon, M. Generalized single excitation configuration interaction: an investigation into the impact of the inclusion of non-orthogonality on the calculation of core-excited states. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8182–8192.

(497) Cederbaum, L.; Tarantelli, F.; Sgamellotti, A.; Schirmer, J. On double vacancies in the core. *J. Chem. Phys.* **1986**, *85*, 6513–6523.

(498) Ambroise, M. A.; Jensen, F. Probing Basis Set Requirements for Calculating Core Ionization and Core Excitation Spectroscopy by the Δ Self-Consistent-Field Approach. *J. Chem. Theory Comput.* **2018**, *15*, 325–337.

(499) Kera, S. et al. High-resolution inner-shell excitation spectroscopy of H 2-phthalocyanine. *J. Chem. Phys.* **2006**, *125*, 014705.

(500) Hayes, W.; Brown, F. C. Absorption by Some Molecular Gases in the Extreme Ultraviolet. *Phys. Rev. A* **1972**, *6*, 21.

(501) Bozek, J.; Tan, K.; Bancroft, G.; Tse, J. High resolution gas phase photoabsorption spectra of SiCl4 and Si (CH3) 4 at the silicon l edges: characterization and assignment of resonances. *Chem. Phys. Lett.* **1987**, *138*, 33–42.

(502) Friedrich, H.; Sonntag, B.; Pittel, B.; Rabe, P.; Schwarz, W. Overlapping core to valence and core to Rydberg transitions and resonances in the XUV spectra of SiF4. *J. Phys. B: At. Mol. Phys.* **1980**, *13*, 25.

(503) Sutherland, D.; Kasrai, M.; Bancroft, G.; Liu, Z.; Tan, K. Si L-and K-edge x-ray-absorption near-edge spectroscopy of gas-phase Si (CH 3) x (OCH 3) 4- x: Models for solid-state analogs. *Phys. Rev. B* **1993**, *48*, 14989.

(504) Liu, Z.; Cutler, J.; Bancroft, G.; Tan, K.; Cavell, R.; Tse, J. High resolution gas phase photoabsorption spectra and multiple-scattering Xα study of PX3 (X= H, CH3, CF3) compounds at the P L2, 3 edge. *Chem. Phys. Lett.* **1990**, *172*, 421–429.

(505) Neville, J. J.; Jürgensen, A.; Cavell, R.; Kosugi, N.; Hitchcock, A. Inner-shell excitation of PF3, PCl3, PCl2CF3, OPF3 and SPF3: Part I. Spectroscopy. *Chem. Phys.* **1998**, *238*, 201–220.

(506) Hu, Y.; Zuin, L.; Püttner, R. High-resolution gas phase P L-edge photoabsorption spectra of PF5. *Can. J. Chem.* **2007**, *85*, 690–694.

(507) Guillemin, R.; Stolte, W.; Dang, L.; Yu, S.-W.; Lindle, D. Fragmentation dynamics of H 2 S following S 2 p photoexcitation. *J. Chem. Phys.* **2005**, *122*, 094318.

(508) Hudson, E. et al. High-resolution measurements of near-edge resonances in the core-level photoionization spectra of SF 6. *Phys. Rev. A* **1993**, *47*, 361.

(509) Hedin, L.; Eland, J.; Karlsson, L.; Feifel, R. An x-ray absorption and a normal Auger study of the fine structure in the S2p- 1 region of the CS2 molecule. *J. Phys. B: At. Mol. Opt. Phys.* **2009**, *42*, 085102.

(510) Ankerhold, U.; Esser, B.; Von Busch, F. Ionization and fragmentation of OCS and CS2 after photoexcitation around the sulfur 2p edge. *Chem. Phys.* **1997**, *220*, 393–407.

(511) Lablanquie, P. et al. Evidence of Single-Photon Two-Site Core Double Ionization of C_2H_2 Molecules. *Phys. Rev. Lett.* **2011**, *107*, 193004.

(512) Aksela, H. et al. Decay channels of core-excited HCl. *Phys. Rev. A* **1990**, *41*, 6000.

(513) Nayandin, O. et al. Angle-resolved two-dimensional mapping of electron emission from the inner-shell 2 p excitations in Cl 2. *Phys. Rev. A* **2001**, *63*, 062719.

(514) Sze, K.; Brion, C. Inner shell and valence shell electronic excitation of ClF3 by high energy electron impact. An investigation of potential barrier effects. *Chem. Phys.* **1989**, *137*, 353–367.

(515) Lu, K.; Chen, J.; Lee, J.; Chen, C.; Chou, T.; Chen, H. State-specific dissociation enhancement of ionic and excited neutral photofragments of gaseous CCl4 and solid-state analogs following Cl 2p core-level excitation. *New J. Phys* **2008**, *10*, 053009.

(516) Hitchcock, A. et al. Inner shell excitation and ionization of the monohalobenzenes. *J. Electron Spectrosc. Relat. Phenom.* **1978**, *13*, 345–360.

(517) Walsh, N. et al. Molecular dynamics of NH 3 induced by core-electron excitation. *Phys. Chem. Chem. Phys.* **2015**, *17*, 18944–18952.

(518) Chen, C.; Ma, Y.; Sette, F. K-shell photoabsorption of the N 2 molecule. *Phys. Rev. A* **1989**, *40*, 6737.

(519) Nakano, M. et al. Single Photon K^{-2} and K^{-1} K^{-1} Double Core Ionization in C_2H_{2n} (n= 1–3), CO, and N_2 as a Potential New Tool for Chemical Analysis. *Phys. Rev. Lett.* **2013**, *110*, 163001.

(520) Salén, P. et al. Experimental verification of the chemical sensitivity of two-site double core-hole states formed by an x-ray free-electron laser. *Phys. Rev. Lett.* **2012**, *108*, 153003.

(521) Barrie, A.; Drummond, I.; Herd, Q. Correlation of calculated and measured 2p spin-orbit splitting by electron spectroscopy using monochromatic x-radiation. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *5*, 217–225.

(522) Chantzis, A.; Kowalska, J. K.; Maganas, D.; DeBeer, S.; Neese, F. Ab initio wave function-based determination of element specific shifts for the efficient calculation of x-ray absorption spectra of main group elements and first row transition metals. *J. Chem. Theory Comput.* **2018**, *14*, 3686–3702.

(523) Lestrange, P. J.; Nguyen, P. D.; Li, X. Calibration of energy-specific TDDFT for modeling K-edge XAS spectra of light elements. *J. Chem. Theory Comput.* **2015**, *11*, 2994–2999.

(524) Besley, N. A. Density Functional Theory Based Methods for the Calculation of X-ray Spectroscopy. *Acc. Chem. Res.* **2020**.

(525) Besley, N. A.; Robinson, D. Theoretical simulation of the spectroscopy and dynamics of a red copper protein. *Faraday Discuss.* **2011**, *148*, 55–70.

(526) Robinson, D.; Besley, N. A. Modelling the spectroscopy and dynamics of plastocyanin. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9667–9676.

(527) Fogarty, R. M. et al. Experimental validation of calculated atomic charges in ionic liquids. *J. Chem. Phys.* **2018**, *148*, 193817.

(528) Buckley, M. W.; Besley, N. A. A theoretical study of the near edge X-ray absorption fine structure of amino acids and proteins. *Chem. Phys. Lett.* **2011**, *501*, 540–546.

(529) Ljubić, I.; Kivimäki, A.; Coreno, M. An experimental NEXAFS and computational TDDFT and ΔDFT study of the gas-phase core excitation spectra of nitroxide free radical TEMPO and its analogues. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10207–10217.

(530) Ljubić, I.; Kivimäki, A.; Coreno, M.; Kazazić, S.; Novak, I. Characterisation of the electronic structure of galvinoxyl free radical by variable energy UPS, XPS and NEXAFS spectroscopy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2480–2491.

(531) Sekino, H.; Bartlett, R. J. A linear response, coupled-cluster theory for excitation energy. *Int. J. Quantum Chem* **1984**, *26*, 255–265.

(532) Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. Energy-specific equation-of-motion coupled-cluster methods for high-energy excited states: Application to K-edge X-ray absorption spectroscopy. *J. Chem. Theory Comput.* **2015**, *11*, 4146–4153.

(533) Coriani, S.; Koch, H. Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework. *J. Chem. Phys.* **2015**, *143*, 181103–181103.

(534) Coriani, S.; Pawłowski, F.; Olsen, J.; Jørgensen, P. Molecular response properties in equation of motion coupled cluster theory: A time-dependent perspective. *J. Chem. Phys.* **2016**, *144*, 024102.

(535) Carbone, J. P.; Cheng, L.; Myhre, R. H.; Matthews, D.; Koch, H.; Coriani, S. In *Advances in Quantum Chemistry*; Elsevier: 2019; Vol. 79, pp 241–261.

(536) Vidal, M. L.; Feng, X.; Epifanovsky, E.; Krylov, A. I.; Coriani, S. New and efficient equation-of-motion coupled-cluster framework for core-excited and core-ionized states. *J. Chem. Theory Comput.* **2019**, *15*, 3117–3133.

(537) Tsuru, S.; Vidal, M. L.; Pápai, M.; Krylov, A. I.; Møller, K. B.; Coriani, S. Time-resolved near-edge X-ray absorption fine structure of pyrazine from electronic structure and nuclear wave packet dynamics simulations. *J. Chem. Phys.* **2019**, *151*, 124114.

(538) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-body theory of core holes. *Phys. Rev. A* **1980**, *22*, 206.

(539) Vidal, M. L.; Pokhilko, P.; Krylov, A. I.; Coriani, S. Equation-of-motion coupled-cluster theory to model L-edge x-ray absorption and photoelectron spectra. *J. Chem. Phys. Lett.* **2020**, *11*, 8314–8321.

(540) Vidal, M. L.; Krylov, A. I.; Coriani, S. Dyson orbitals within the fc-CVS-EOM-CCSD framework: theory and application to X-ray photoelectron spectroscopy of ground and excited states. *Phys. Chem. Chem. Phys.* **2020**, *22*, 2693–2703.

(541) Norman, P.; Dreuw, A. Simulating X-ray spectroscopies and calculating core-excited states of molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.

(542) List, N. H.; Dempwolff, A. L.; Dreuw, A.; Norman, P.; Martınez, T. J. Probing competing relaxation pathways in malonaldehyde with transient X-ray absorption spectroscopy. *Chem. Sci.* **2020**, *11*, 4180–4193.

(543) Wenzel, J.; Dreuw, A. Physical properties, exciton analysis, and visualization of core-excited states: An intermediate state representation approach. *J. Chem. Theory Comput.* **2016**, *12*, 1314–1330.

(544) Wenzel, J.; Holzer, A.; Wormit, M.; Dreuw, A. Analysis and comparison of CVS-ADC approaches up to third order for the calculation of core-excited states. *J. Chem. Phys.* **2015**, *142*, 214104.

(545) Myhre, R. H.; Coriani, S.; Koch, H. Near-edge X-ray absorption fine structure within multilevel coupled cluster theory. *J. Chem. Theory Comput.* **2016**, *12*, 2633–2643.

(546) Hait, D.; Head-Gordon, M. Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining sub eV Error from a Restricted Open-Shell Kohn-Sham Approach. *J. Phys. Chem. Lett.* **2020**, *11*, 775–786.

(547) Ehlert, C.; Klamroth, T. PSIXAS: A Psi4 plugin for efficient simulations of X-ray absorption spectra based on the transition-potential and Δ-Kohn-Sham method. *J. Comp. Chem.* **2020**, *41*, 1781–1789.

(548) Yang, Z. et al. Electron-Withdrawing Effects in the Photodissociation of CH2ICl To Form CH2Cl Radical, Simultaneously Viewed Through the Carbon K and Chlorine L2, 3 X-ray Edges. *J. Am. Chem. Soc.* **2018**, *140*, 13360–13366.

(549) Yamaguchi, K. Singlet unrestricted Hartree-Fock Slater (UHFS) model for unstable metal-metal bonds. *Chem. Phys. Lett.* **1979**, *66*, 395–399.

(550) Noodleman, L. Valence bond description of antiferromagnetic coupling in transition metal dimers. *J. Chem. Phys.* **1981**, *74*, 5737–5743.

(551) Noodleman, L.; Norman Jr, J. G.; Osborne, J. H.; Aizman, A.; Case, D. A. Models for ferredoxins: electronic structures of iron-sulfur clusters with one, two, and four iron atoms. *J. Am. Chem. Soc.* **1985**, *107*, 3418–3426.

(552) Sinnecker, S.; Neese, F.; Noodleman, L.; Lubitz, W. Calculating the electron paramagnetic resonance parameters of exchange coupled transition metal complexes using broken symmetry density functional theory: application to a MnIII/MnIV model compound. *J. Am. Chem. Soc.* **2004**, *126*, 2613–2622.

(553) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. FeMo cofactor of nitrogenase: a density functional study of states MN, MOX, MR, and MI. *J. Am. Chem. Soc.* **2001**, *123*, 12392–12410.

(554) Adams, D. M.; Noodleman, L.; Hendrickson, D. N. Density Functional Study of the Valence-Tautomeric Interconversion Low-Spin [CoIII (SQ)(Cat)(phen)] $\rightleftharpoons$ High-Spin [CoII (SQ) 2 (phen)]. *Inorg. Chem.* **1997**, *36*, 3966–3984.

(555) Mouesca, J.-M.; Noodleman, L.; Case, D. A. Density-functional calculations of spin coupling in [Fe4S4] 3+ clusters. *Int. J. Quantum Chem* **1995**, *56*, 95–102.

(556) Witzke, R. J.; Hait, D.; Chakarawet, K.; Head-Gordon, M.; Tilley, T. D. Bimetallic Mechanism for Alkyne Cyclotrimerization with a Two-Coordinate Fe Precatalyst. *ACS Catalysis* **2020**.

(557) Yang, Z. Characterization of Substituted Radicals by Multi-Edge Femtosecond X-Ray Transient Absorption Spectroscopy, Ph.D. Thesis, UC Berkeley, 2018.

(558) Alagia, M. et al. Probing the potential energy surface by high-resolution x-ray absorption spectroscopy: The umbrella motion of the core-excited C H 3 free radical. *Phys. Rev. A* **2007**, *76*, 022509.

(559) Alagia, M. et al. The soft X-ray absorption spectrum of the allyl free radical. *Phys. Chem. Chem. Phys.* **2013**, *15*, 1310–1318.

(560) Couto, R. C. et al. The carbon and oxygen K-edge NEXAFS spectra of CO+. *Phys. Chem. Chem. Phys.* **2020**.

(561) Parent, P. et al. The irradiation of ammonia ice studied by near edge x-ray absorption spectroscopy. *J. Chem. Phys.* **2009**, *131*, 154308.

(562) Lindblad, R. et al. X-Ray Absorption Spectrum of the N 2+ Molecular Ion. *Phys. Rev. Lett.* **2020**, *124*, 203001.

(563) Bari, S. et al. Inner-shell X-ray absorption spectra of the cationic series NH y+(y= 0–3). *Phys. Chem. Chem. Phys.* **2019**, *21*, 16505–16514.

(564) Zhang, W.; Sze, K.; Brion, C.; Tong, X.; Li, J. Inner-shell electron energy loss spectra of NO2 at high resolution: Comparison with multichannel quantum defect calculations of dipole oscillator strengths and transition energies. *Chem. Phys.* **1990**, *140*, 265–279.

(565) Stranges, S.; Richter, R.; Alagia, M. High-resolution inner-shell photoabsorption of the OH and OD free radicals. *J. Chem. Phys.* **2002**, *116*, 3676–3680.

(566) Lacombe, S.; Bournel, F.; Laffon, C.; Parent, P. Radical Photochemistry in Oxygen-Loaded Ices. *Angew. Chem. Int. Ed.* **2006**, *45*, 4159–4163.

(567) Coreno, M. et al. Vibrationally resolved oxygen K→ π^* spectra of O2 and CO. *Chem. Phys. Lett.* **1999**, *306*, 269–274.

(568) Gill, P. M.; Pople, J. A.; Radom, L.; Nobes, R. H. Why does unrestricted Mo/ller–Plesset perturbation theory converge so slowly for spin-contaminated wave functions? *J. Chem. Phys.* **1988**, *89*, 7307–7314.

(569) Johansson, K.; Head-Gordon, M.; Schrader, P.; Wilson, K.; Michelsen, H. Resonance-stabilized hydrocarbon-radical chain reactions may explain soot inception and growth. *Science* **2018**, *361*, 997–1000.

(570) Woon, D. E.; Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

(571) Burton, H. G.; Wales, D. J. Energy Landscapes for Electronic Structure. *J. Chem. Theory Comput.* **2021**, *17*, 151–169.

(572) Levi, G.; Ivanov, A. V.; Jónsson, H. Variational density functional calculations of excited states via direct optimization. *J. Chem. Theory. Comput.* **2020**, *16*, 6968–6982.

(573) Carter-Fenk, K.; Herbert, J. M. State-Targeted Energy Projection: A Simple and Robust Approach to Orbital Relaxation of Non-Aufbau Self-Consistent Field Solutions. *J. Chem. Theory Comput.* **2020**, *16*, 5067–5082.

(574) Grofe, A. et al. Generalization of block-localized wave function for constrained optimization of excited determinants. *J. Chem. Theory Comput.* **2020**, *17*, 277–289.

(575) Corzo, H. H.; Abou Taka, A.; Pribram-Jones, A.; Hratchian, H. P. Using projection operators with maximum overlap methods to simplify challenging self-consistent field optimization. *J. Comp. Chem.* **2022**, *43*, 382–390.

(576) Ambroise, M. A.; Dreuw, A.; Jensen, F. Probing basis set requirements for calculating core ionization and core excitation spectra using correlated wave function methods. *J. Chem. Theory Comput.* **2021**, *17*, 2832–2842.

(577) DiRisio, R. J.; Jones, C. M.; Ma, H.; Rousseau, B. J. Viewpoints on the 2020 Virtual Conference on Theoretical Chemistry, 2020.

(578) Bussy, A.; Hutter, J. First-principles correction scheme for linear-response time-dependent density functional theory calculations of core electronic states. *J. Chem. Phys.* **2021**, *155*, 034108.

(579) Kunze, L.; Hansen, A.; Grimme, S.; Mewes, J.-M. PCM-ROKS for the Description of Charge-Transfer States in Solution: Singlet–Triplet Gaps with Chemical Accuracy from Open-Shell Kohn–Sham Reaction-Field Calculations. *J. Phys. Chem. Lett.* **2021**, *12*, 8470–8480.

(580) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *Chem. Phys. Chem.* **2011**, *12*, 3077–3094.

(581) Cunha, L. A.; Hait, D.; Kang, R.; Mao, Y.; Head-Gordon, M. Relativistic Orbital-Optimized Density Functional Theory for Accurate Core-Level Spectroscopy. *J. Chem. Phys. Lett.* **2022**, *13*, 3438–3449.

(582) Hait, D.; Oosterbaan, K. J.; Carter-Fenk, K.; Head-Gordon, M. Computing X-ray absorption spectra from linear-response particles atop optimized holes. *arXiv preprint arXiv:2203.13529* **2022**.

(583) Burton, H. G. Energy Landscape of State-Specific Electronic Structure Theory. *J. Chem. Theory. Comput.* **2022**, *18*, 1512–1526.

(584) Kossoski, F.; Marie, A.; Scemama, A.; Caffarel, M.; Loos, P.-F. Excited States from State-Specific Orbital-Optimized Pair Coupled Cluster. *J. Chem. Theory Comput.* **2021**, *17*, 4756–4768.

(585) Arias-Martinez, J. E.; Cunha, L. A.; Oosterbaan, K. J.; Lee, J.; Head-Gordon, M. Core excitations: how far can we push correlated single-reference formalism for Delta-based methods? *arXiv preprint arXiv:2204.13056* **2022**.

(586) Hanscam, R.; Neuscamman, E. Applying generalized variational principles to excited-state-specific complete active space self-consistent field theory. *arXiv preprint arXiv:2111.02590* **2021**.