

On Improving Density Functional Approximations: When Optimization Matters

by

Zehua Chen

Department of Chemistry
Duke University

Date: _____

Approved:

Weitao Yang, Advisor

David N. Beratan

Jianfeng Lu

Peng Zhang

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Chemistry
in the Graduate School of
Duke University

2021

ABSTRACT

On Improving Density Functional Approximations: When
Optimization Matters

by

Zehua Chen

Department of Chemistry
Duke University

Date: _____

Approved: _____

Weitao Yang, Advisor

David N. Beratan

Jianfeng Lu

Peng Zhang

An abstract of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Chemistry
in the Graduate School of
Duke University

2021

Copyright © 2021 by Zehua Chen
All rights reserved

Abstract

The Kohn-Sham density functional theory has been the most popular method in electronic structure calculations. Although the theory is in principle exact, approximations are needed for the exchange-correlation energy to make practical calculations possible. This dissertation focuses on one aspect to seek for better approximations: optimization. Optimization has many different meanings in quantum chemistry. Specifically, we carry out energy optimization for a few known energy forms to get insight about what self-consistency can achieve. One kind of them is designed from linear response theories (two-electron addition energies from the particle-particle random phase approximation and spin-flip excitation energies from the spin-flip linear-response time-dependent density functional theory). The post-self-consistency versions are already known to produce excellent results for states that involve static correlations, while the energy optimization further improves the accuracy, especially when the total energy is a quantity of interest. The optimization technique used is the generalized optimized effective potential method, which is a non-local counterpart to the well-established optimized effective potential method. The other part is related to the local orbital scaling correction method, which mostly deals with delocalization errors. We designed a robust self-consistent workflow to avoid the difficult exact-Hamiltonian derivation, and observed reliable prediction on electron densities, total energies, and energy-level alignments. Machine-learning can also be considered as an optimization technique, which is used to minimize the errors of a density functional approximation in the provided database. Special treatment was carried out to transfer electron densities in 3D to neural network compatible input variables (translational and rotational invariant; discrete numbers). The outcome is a finite-range non-local density functional correction to low cost functionals, which

is able to achieve similar level of accuracy as compared to accurate and expensive functionals, while keeping the computational cost low. In summary, “optimization” over existing functional approximations is a reliable way to get an improvement, and sometimes new physics can be discovered. On the other hand, “optimization” can also be done by “machines”, which may become more and more popular in the future.

Acknowledgements

First of all, I want to thank my advisor, Prof. Weitao Yang. He is a great scientist who always has many interesting ideas. I learned from him to keep motivated about research projects. At those hard times when the progress got stuck, or months of efforts got little meaningful pay back, he is the one who got me the solution.

Next, I want to thank previous collaborators, who are Yang Yang, Du Zhang, Neil Qiang Su and Ye Jin. We had a lot of wonderful discussions and they taught me a lot. I also received tremendous amount of help from them.

I want to say a big “Thank you!” to Yuncai Mei, who is actively collaborating with me now. He is such a hardworking person that often keeps me moving. Without his help my Ph.D. journey would have been a lot harder.

I also want to thank Prof. Patrick Charbonneau, who was previously my preliminary exam committee member. I like his class (and was definitely interested in statistical mechanics at that time), and I got his care about my future last year.

My thanks also go to Jiachen Li and Pan Zhang. Jiachen is a professional in many-body theories that I need to learn from. Pan had quite some discussions with me about machine learning, which indeed helped.

Friends can make hard Ph.D. life easier and happier. Xuyan Ru, Lianjun Zheng and He Meng, though many of them have already left Duke way earlier than me, will never be forgotten. My gaming friends too, Zhongxing Xu, Zhuoran Long and Haomin Wang, who seldom have the chance to meet in real life, but regularly on the Internet.

Last but not the least, I want to thank my parents. I have gotten so much love and care from them, which greatly supported me to carry on the Ph.D. journey.

Contents

Abstract	iv
Acknowledgements	vi
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 From the Schrödinger equation to the density functional theory	1
1.1.1 Challenges in making accurate density functional approximations	2
1.2 The generalized optimized effective potential (GOEP) method	3
2 Potential Optimization of a Density Functional Designed for Static Correlation	6
2.1 Introduction	6
2.2 Methods	7
2.2.1 The physical density from the reference density matrix	10
2.2.2 Optimization procedure	12
2.2.3 Analytic gradients for the pp-RPA and the SF-TDA	16
2.3 Results	17
3 Machine-Learn a Density Functional for Optimal Accuracy	25
3.1 Introduction	25
3.2 Many-body Expansion in the DFT	27
3.3 Machine-Learn the Finite-Range Non-Local Functional	31
3.4 Methods	32

3.4.1	Details about the neural network	35
3.4.2	Database rescaling	36
3.5	Computational Details	36
3.6	Results and Discussion	39
4	Improved Self-consistency for a Functional Tackles Delocalization Errors	52
4.1	Introduction	52
4.2	Methods	56
4.2.1	Macro-SCF-LOSC	61
4.3	Computational Details	62
4.4	Results	63
5	Conclusions	75
	Bibliography	77
	Biography	94

List of Figures

2.1	Relative energies of H ₂ , LiH, BH and HF molecules with respect to corresponding dissociated atoms.	19
2.2	Relative energy of HF molecule with respect to dissociated atoms. . .	21
2.3	Relative energy of twisted ethylene with respect to flat ethylene. . . .	22
3.1	The electron density along the chemical bond of the HF molecule. . .	28
3.2	Comparison of different exchange-correlation energy expansion schemes	30
3.3	An example plot of the radial functions used when generating the power spectrum.	35
3.4	The number of times when an element appears in the molecules that are used by the database.	38
3.5	Superposition of the power spectrum computed on the entire database.	40
3.6	Cross-validation errors for different neural network parameters.	41
3.7	Cross-validation errors for a few different setups of μ_{\max} and c , with BLYP base functional.	42
3.8	Error plot of the BLYP-based neural network functional, along with other methods.	45
3.9	Error plot of the PBE-based neural network functional, along with other methods.	46
3.10	Fractional charge error of the carbon atom.	50
4.1	Original SCF procedure with the approximate LOSC effective Hamiltonian.	58
4.2	New SCF procedure with the exact LOSC Hamiltonian.	60
4.3	The macro-SCF procedure for LOSC.	61

4.4	Comparison of SCF performance between SCF-LOSC2 and the original SCF-LOSC2 with an approximate Hamiltonian for polyacetylene. . .	64
4.5	Dissociation of the LiF molecule.	69
4.6	Dissociation of the LiH molecule.	70
4.7	Dissociation of the HF molecule.	71
4.8	Dissociation of the donor-acceptor complex.	72
4.9	HOMO and LUMO energy-level alignment of the donor-acceptor complex.	73

List of Tables

2.1	Binding energies (eV) for H ₂ , LiH, BH and HF.	20
2.2	Equilibrium bond lengths (Å) for H ₂ , LiH, BH and HF.	20
2.3	Atomic excitation energies (eV).	23
2.4	Molecular vertical excitation energies (eV).	24
3.1	Comparing neural network functionals with ω B97M-V for training data belong to MGCDB84.	43
3.2	Comparing neural network functionals with DSD-BLYP-D3(BJ) for training data belong to GMTKN55.	43
3.3	Comparing neural network functionals with ω B97M-V for tests belong to MGCDB84.	44
3.4	Comparing neural network functionals with DSD-BLYP-D3(BJ) for tests belong to GMTKN55.	50
3.5	Errors of 30 non-covalent interaction energies in the HAL59 set, and 28 non-covalent interaction energies in the HEAVY28 set.	50
3.6	Errors in the MOR41 set.	51
4.1	Testing on the F atom for the effect of the orientation of initial guess to the new SCF-LOSC approach.	65
4.2	Total energies comparison between SCF-LOSC2 and macro-SCF-LOSC2.	66
4.3	Mulliken charge comparison between SCF-LOSC2 and macro-SCF-LOSC2.	66
4.4	Mean absolute errors (MAEs) for atomization energies (AE), reaction barriers (RB), IPs and EAs from different methods.	67
4.5	Negative HOMO energies of polyacetylene with different unit numbers.	67

Chapter 1

Introduction

1.1 From the Schrödinger equation to the density functional theory

In the non-relativistic limit, the solution to the Schrödinger equation fully determines the property of a system at any time t :

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}, t) = \hat{H} \Psi(\mathbf{x}, t), \quad (1.1)$$

which usually requires the solution to an eigenvalue problem:

$$\hat{H} \Psi(\mathbf{x}) = E \Psi(\mathbf{x}). \quad (1.2)$$

However, obtaining solutions to above equations are always challenging and therefore require approximations. After the Born-Oppenheimer approximation, the Hamiltonian has electron's quantum degree of freedom left only:

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_A \sum_i \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (1.3)$$

where nuclei are treated as fixing dots that only produce Coulomb attraction forces to electrons. This yields a popular field in quantum chemistry, which mostly works on the electronic structure theory.

Using the linear combination of Slater determinants as an approximation to the

wavefunction has led to a huge branch of the electronic structure theory: the wavefunction theory. With the full-configuration interaction being the fundamentally exact method and couple-cluster methods being the golden standard in quantum chemistry, the wavefunction theory definitely has won many applications. However, with the rapid growth of research interest in simulating large molecules and bulk systems, the huge computational cost of wavefunction methods makes them not the best choice.

Unlike the wavefunction theory, the density functional theory (DFT) [1, 2, 3] is formally exact with the need of the electron density only. This tremendously speeds up the calculation. The past decades have witnessed the unstoppable trend of using DFT more and more in quantum chemistry simulations. However, the approximation to the unknown exchange-correlation functional may not be as accurate as wavefunction methods, and conventional approximations usually suffer from a number of defects.

1.1.1 Challenges in making accurate density functional approximations

Designing an accurate density functional approximation is nearly as hard as solving the many-body problem exactly. Currently available approximations usually suffer from delocalization errors, [4, 5, 6, 7] static correlation errors, and are not accurate enough even when these errors are absent. To overcome the difficulties, this dissertation has three parts to work on:

1. Static correlation errors, by designing an approximation with multireference included;
2. Delocalization errors, by improving an existing method that works on the de-

localization error;

3. Accuracy issues, by optimizing a neural network functional with respect to data sets.

The key to improving density functional approximations in this dissertation is “optimization”. With or without optimization, a method may exhibit quite different behaviors.

1.2 The generalized optimized effective potential (GOEP) method

Since the generalized optimized effective potential (GOEP) method will be used, we review the method here. The idea of the GOEP method is similar to the direct optimization version of local optimized effective potential method.[8, 9] Unlike the local OEP method, the GOEP method does not need to be supplied with a base potential or additional basis functions. The non-local potential is directly expressed in terms of matrix elements in the molecular orbital space.

The main purpose of the GOEP method is to optimize the nonlocal effective potential in the generalized Kohn-Sham scheme instead of directly optimizing the orbitals or ρ_s^{ref} , because the effective potential determines the density matrix ρ_s^{ref} . The chain rule can thus be applied:

$$\frac{\delta E}{\delta \rho_s^{\text{ref}}(\mathbf{x}', \mathbf{x})} = \iint d\mathbf{y} d\mathbf{y}' \frac{\delta v^{\text{GOEP}}(\mathbf{y}, \mathbf{y}')}{\delta \rho_s^{\text{ref}}(\mathbf{x}', \mathbf{x})} \frac{\delta E}{\delta v^{\text{GOEP}}(\mathbf{y}, \mathbf{y}')} \quad (1.4)$$

Hence, the stationary condition is met in the nonlocal effective potential optimization, the total energy is also stationary with respect to the variation of the reference density

matrix:

$$\frac{\delta E}{\delta v^{\text{GOEP}}} = 0 \rightarrow \frac{\delta E}{\delta \rho_s^{\text{ref}}} = 0. \quad (1.5)$$

Given a set of generalized Kohn-Sham orbitals $\{|\phi_p\rangle\}$, a small variation in the effective potential δv^{GOEP} changes in the orbitals. According to the perturbation theory, the first-order orbital variation is

$$|\delta\phi_p\rangle = -\sum_{q \neq p} |\phi_q\rangle \frac{\langle \phi_q | \delta v^{\text{GOEP}} | \phi_p \rangle}{\epsilon_q - \epsilon_p} = -\sum_{q \neq p} |\phi_q\rangle \frac{(\delta v^{\text{GOEP}})_{qp}}{\epsilon_q - \epsilon_p}. \quad (1.6)$$

With $p \neq q$, we have

$$\frac{\delta\phi_r(\mathbf{x})}{(\delta v^{\text{GOEP}})_{qp}} = -\delta_{pr} \frac{\phi_q(\mathbf{x})}{\epsilon_q - \epsilon_p} = \frac{\delta\phi_r(\mathbf{x})}{(\delta v^{\text{GOEP}})_{pq}^*} = \left(\frac{\delta\phi_r^*(\mathbf{x})}{(\delta v^{\text{GOEP}})_{pq}} \right)^*. \quad (1.7)$$

Thus the potential derivative of a pure orbital dependent total energy functional is

$$\frac{\delta E}{(\delta v^{\text{GOEP}})_{qp}} = \sum_r \int d\mathbf{x} \left[\frac{\delta E}{\delta\phi_r^*(\mathbf{x})} \frac{\delta\phi_r^*(\mathbf{x})}{(\delta v^{\text{GOEP}})_{qp}} + \frac{\delta E}{\delta\phi_r(\mathbf{x})} \frac{\delta\phi_r(\mathbf{x})}{(\delta v^{\text{GOEP}})_{qp}} \right] \quad (1.8)$$

$$= -\sum_r \int d\mathbf{x} \left[\frac{\delta E}{\delta\phi_r^*(\mathbf{x})} \frac{\delta_{qr}\phi_p^*(\mathbf{x})}{\epsilon_p - \epsilon_q} + \frac{\delta E}{\delta\phi_r(\mathbf{x})} \frac{\delta_{pr}\phi_q(\mathbf{x})}{\epsilon_q - \epsilon_p} \right] \quad (1.9)$$

$$= \frac{1}{\epsilon_q - \epsilon_p} \int d\mathbf{x} \left[\phi_p^*(\mathbf{x}) \frac{\delta E}{\delta\phi_q^*(\mathbf{x})} - \frac{\delta E}{\delta\phi_p(\mathbf{x})} \phi_q(\mathbf{x}) \right] \quad (1.10)$$

$$= \left(\frac{\delta E}{(\delta v^{\text{GOEP}})_{pq}} \right)^*. \quad (1.11)$$

The derivatives of the reference energy can then be derived, with only occupied-virtual and virtual-occupied blocks having non-zero values, thanks to the unitary invariance of KS-DFT/Hartree-Fock:

$$\frac{\delta E^{\text{ref}}[\rho_s^{\text{ref}}]}{\delta (v^{\text{GOEP}})_{ai}} = -\frac{\langle \phi_i | \hat{F} | \phi_a \rangle}{\epsilon_a - \epsilon_i} = \left(\frac{\delta E^{\text{ref}}[\rho_s^{\text{ref}}]}{\delta (v^{\text{GOEP}})_{ia}} \right)^*, \quad (1.12)$$

where \hat{F} is the Fock operator constructed with ρ_s^{ref} . The gradient matrix is symmetric for real orbitals.

During the optimization procedure, actually the generalized Kohn-Sham one-particle Hamiltonian matrix is stored and updated instead of the potential. Because $\delta H_s = \delta v^{\text{GOEP}}$, one simply has

$$\frac{\delta E}{(\delta H_s)_{qp}} = \frac{\delta E}{(\delta v^{\text{GOEP}})_{qp}}. \quad (1.13)$$

The detailed procedure for the optimization is:

1. Start from a set of orbitals, use DFA orbital energies as the initial guess for GOEP eigenvalues. These eigenvalues can be freely adjusted, as long as the orbital ordering is preserved. Initialize the H_s matrix.
2. Compute the gradients.
3. Update the H_s matrix according to the optimization algorithm.
4. Get a new set of orbitals $\{\phi_p\}$ from the diagonalization of H_s .
5. With $\{\phi_p\}$, compute the reference system DFA energy and do the pp-RPA or SF-TDDFT calculation.
6. Compute the gradients. If the convergence criteria is not met, go back to step 3.

Chapter 2

Potential Optimization of a Density Functional Designed for Static Correlation

This chapter is mainly adapted from the journal article:

Zehua Chen, Du Zhang, Ye Jin, Yang Yang, Neil Qiang Su, and Weitao Yang. “Multireference density functional theory with generalized auxiliary systems for ground and excited states.” *The Journal of Physical Chemistry Letters*, **2017**, 8(18), 4479-4485.

2.1 Introduction

The past half-century has seen increasingly rapid advances in the density functional theory (DFT) [1, 2, 3, 10]. While the Kohn-Sham DFT (KS DFT), is exact in principle, approximations to the exchange-correlation functional are necessary for practical calculations[1, 2, 3]. Many approximate exchange-correlation energy functionals have been proposed, such as local density approximations (LDAs) [11, 12], generalized gradient approximations (GGAs) [13, 14, 15, 16] and hybrid functionals [17, 18]. These common density functional approximations (DFA), have been extensively applied in chemistry, solid state physics and biology.

However, there are still many cases in which conventional approximations fail. Bond breaking is an example where single determinantal methods such as restricted Hartree-Fock and KS DFT fail. It has been shown that errors in the dissociation energy can be decomposed into fractional charge errors and fractional spin errors[19].

Particularly, the fractional spin error occurs when the system is multiconfigurational. In this case, the static correlation, which is included by multireference methods, has significant contributions to the total energy. In the wave function theory (WFT), multiconfiguration self-consistent field (MCSCF) methods, the complete active space self-consistent field (CASSCF) method[20] and other multireference methods, are used to properly describe near-degenerate states. Dynamic correlation is further taken into account by post-SCF methods. Multireference perturbation theories, such as the complete-active-space second-order perturbation theory (CASPT2)[21], or multireference configuration interactions (MRCI)[22], are typical post-SCF methods using the MCSCF wave function as the starting point. Attempts have also been made on the development of multireference DFT. For example, a range separated scheme, which treats the long range part by MRCI and the short range part by DFT, is available[23]. There are some other approaches that combine DFT calculations with multiconfiguration WFT. For example, in the multistate density functional theory (MSDFT), multiple KS DFT calculations are carried out, followed by a configuration interaction (CI)-like calculation to produce ground and excited state energies as a weighted sum[24]. The multiconfiguration pair density functional theory (MC-PDFT)[25] uses the MCSCF wave function to produce the on-top pair density, then evaluates the total energy by approximate functionals.

2.2 Methods

In this chapter, we present a new approach to multireference DFT. This method uses a generalized auxiliary reference system. However, the auxiliary system is not designed to describe the interacting system density. The total energy of the physical system, $E_v[\rho]$, consists of two parts: $E_v^{\text{ref}}[\rho_s^{\text{ref}}]$, the energy of the auxiliary system, which is

determined with a chosen density functional approximation (DFA), and $\Delta E[\rho_s^{\text{ref}}]$, the excitation energy from an approximate linear response theory that restores the symmetry to that of the physical system, which leads to multi-determinant description of the physical system naturally and rigorously. Thus,

$$E_v[\rho] = E_v^{\text{ref}}[\rho_s^{\text{ref}}] + \Delta E[\rho_s^{\text{ref}}], \quad (2.1)$$

where ρ is the density of the physical system to be studied and ρ_s^{ref} is the density matrix of the non-interacting auxiliary reference system. The excited state energy functionals for the physical system are directly available by selecting different target states in the ΔE term:

$$E_{v,n}[\rho_n] = E_v^{\text{ref}}[\rho_s^{\text{ref}}] + \Delta E_n[\rho_s^{\text{ref}}], \quad (2.2)$$

where n indicates the energy level. In combination with the linear response theory, multireference effects are included. Specifically, N -electron physical systems can be recovered from an $(N - 2)$ -electron system with two-electron addition excitation energies provided by the particle-particle random phase approximation (pp-RPA)[26, 27, 28]. Similarly, low-spin states can be generated by spin-flip excitations from the linear response theory as in the spin-flip time-dependent DFT (SF-TDDFT)[29, 30, 31, 32], from a high-spin reference. These two methods are only examples to demonstrate our approach. There may exist other variants such as two-electron removal, one-electron addition/removal and multiple spin flips.

Note that without self-consistently optimizing the electron density of the physical system ρ , Eqs (2.1-2.2) revert to known approaches within the context of the pp-RPA and the SF-TDDFT for ground and excited energy calculations based on the self-

consistent optimization of $E_v^{\text{ref}}[\rho_s^{\text{ref}}]$ with respect to ρ_s^{ref} , which is just a regular ground state (G)KS calculation for the auxiliary system. This simplification recovers post-(G)KS calculations of the ground and excited states of the physical system[29, 31, 33].

The key departure in our work is to recognize that Eqs (2.1-2.2) as total energy functionals of the physical system density ρ . This functional has major differences from the traditional (G)KS functional – it does not have the form of common (G)KS DFT and it naturally and rigorously has multiconfiguration description of the ground and excited states. Viewed as a density functional, the ground state energy is given by the minimum of Eq. (2.1):

$$E_v(N) = \min_{\rho \rightarrow N} E_v[\rho], \quad (2.3)$$

and the energies of excited states are given by the stationary solutions of $E_{v,n}[\rho_n]$. The electron density, $\rho(\mathbf{r})$, of the physical system is different from ρ_s^{ref} , that of the auxiliary system, and is uniquely determined from the functional derivative of the total energy with respect to the external potential ($v(\mathbf{r})$):

$$\rho(\mathbf{r}) = \left(\frac{\delta E_v(N)}{\delta v(\mathbf{r})} \right)_N. \quad (2.4)$$

In order to achieve self-consistency required in Eq. (2.3), optimization with respect to the physical density ρ is necessary, which would seem very challenging. However, since the total energy, Eq. (2.1), is given explicitly in terms of the auxiliary density matrix ρ_s^{ref} , and the physical density ρ is uniquely determined by ρ_s^{ref} as in Eq. (2.4) The following subsection gives a specific example.

2.2.1 The physical density from the reference density matrix

Here we use the SF-CIS (SF-TDA@HF) as an example to show how the physical density is generated by the reference density matrix following the Eq. (4) in the text.

The total energy is

$$E_v = E_v^{\text{ref}} + \Delta E \quad (2.5)$$

$$\begin{aligned} &= \sum_i^N \langle i | \hat{F} | i \rangle - \frac{1}{2} \sum_i^N \sum_j^N \langle ij || ij \rangle \\ &+ \sum_{\bar{a}\bar{b}j} (X_{\bar{a}i})^* \left(\langle \bar{a} | \hat{F} | \bar{b} \rangle \langle i | j \rangle - \langle i | j \rangle \langle \bar{a} | \bar{b} \rangle - \langle i \bar{b} | j \bar{a} \rangle \right) X_{\bar{b}j}. \end{aligned} \quad (2.6)$$

When a perturbation in the external potential v is introduced, the Fock operator changes,

$$\hat{F} \rightarrow \hat{F} + \delta v(\mathbf{r}), \quad (2.7)$$

leading to changes in orbitals:

$$|p\rangle \rightarrow |p\rangle + |\delta p\rangle. \quad (2.8)$$

The full derivative is

$$\rho(\mathbf{r}) = \frac{\delta E_v}{\delta v(\mathbf{r})} \quad (2.9)$$

$$= \sum_p^K \frac{\delta E_v}{\delta \phi_p} \frac{\delta \phi_p}{\delta v} + c.c. + \frac{\partial E_v}{\partial v}. \quad (2.10)$$

This gets greatly simplified as the SCF solution is optimal with respect to the orbitals $\delta E_v / \delta \phi_p = 0$. Assume the orbitals are all real,

$$\rho(\mathbf{r}) = + \frac{\partial E_v}{\partial v} \quad (2.11)$$

$$= \sum_i^N \phi_i^2(\mathbf{r}) + \sum_{\bar{a}\bar{i}\bar{b}\bar{j}} (X_{\bar{a}\bar{i}})^* X_{\bar{b}\bar{j}} [\phi_{\bar{a}}(\mathbf{r})\phi_{\bar{b}}(\mathbf{r})\delta_{\bar{i}\bar{j}} - \phi_{\bar{i}}(\mathbf{r})\phi_{\bar{j}}(\mathbf{r})\delta_{\bar{a}\bar{b}}] \quad (2.12)$$

$$= \rho^{\text{ref}}(\mathbf{r}) + \sum_{\bar{a}\bar{b}\bar{i}} (X_{\bar{a}\bar{i}})^* X_{\bar{b}\bar{i}} \phi_{\bar{a}}(\mathbf{r})\phi_{\bar{b}}(\mathbf{r}) - \sum_{\bar{a}\bar{i}\bar{j}} (X_{\bar{a}\bar{i}})^* X_{\bar{a}\bar{j}} \phi_{\bar{i}}(\mathbf{r})\phi_{\bar{j}}(\mathbf{r}). \quad (2.13)$$

It can be verified that this density is identical to what is produced by SF-CIS, as shown in the following.

The physical system wavefunction can be expressed as

$$|\Phi\rangle = \sum_{\bar{a}\bar{i}} X_{\bar{a}\bar{i}} |\Psi_{\bar{i}}^{\bar{a}}\rangle, \quad (2.14)$$

where $|\Psi_{\bar{i}}^{\bar{a}}\rangle$ is a singly spin-flip excited determinant generated from the high-spin

determinant $|\Psi_0\rangle$.

$$\rho(\mathbf{r}) = N \int ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N \Phi(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N)^* \Phi(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N) \quad (2.15)$$

$$= N \int ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N \sum_{\bar{a}\bar{b}j} (X_{\bar{a}i})^* X_{\bar{b}j} \Psi_i^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N)^* \Psi_j^{\bar{b}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N) \quad (2.16)$$

$$= N \int ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N \sum_{\bar{a}i} (X_{\bar{a}i})^* X_{\bar{a}i} \Psi_i^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N)^* \Psi_i^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N) \\ + \sum_{\bar{a} \neq \bar{b}, i} (X_{\bar{a}i})^* X_{\bar{b}i} \Psi_i^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N)^* \Psi_i^{\bar{b}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N) \\ + \sum_{\bar{a}, i \neq j} (X_{\bar{a}i})^* X_{\bar{a}j} \Psi_i^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N)^* \Psi_j^{\bar{a}}(\mathbf{x}, \mathbf{x}_2, \cdots, \mathbf{x}_N) \quad (2.17)$$

$$= N \left[\sum_{\bar{a}i} (X_{\bar{a}i})^* X_{\bar{a}i} (\rho^{\text{ref}}(\mathbf{r}) + \phi_a^2(\mathbf{r}) - \phi_i^2(\mathbf{r})) + \sum_{\bar{a} \neq \bar{b}, i} (X_{\bar{a}i})^* X_{\bar{b}i} \phi_{\bar{a}}(\mathbf{r}) \phi_{\bar{b}}(\mathbf{r}) \right. \\ \left. - \sum_{\bar{a}, i \neq j} (X_{\bar{a}i})^* X_{\bar{a}j} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}) \right] \times \frac{(N-1)!}{N!} \quad (2.18)$$

$$= \rho^{\text{ref}}(\mathbf{r}) + \sum_{\bar{a}\bar{b}i} (X_{\bar{a}i})^* X_{\bar{b}i} \phi_{\bar{a}}(\mathbf{r}) \phi_{\bar{b}}(\mathbf{r}) - \sum_{\bar{a}ij} (X_{\bar{a}i})^* X_{\bar{a}j} \phi_i(\mathbf{r}) \phi_j(\mathbf{r}). \quad (2.19)$$

2.2.2 Optimization procedure

We can view the energy also as a functional of ρ_s^{ref} explicitly:

$$E_v[\rho] = E_v[\rho_s^{\text{ref}}[\rho]]. \quad (2.20)$$

Consequently, the variation with respect to ρ is directly connected to the variation with respect to ρ_s^{ref} :

$$\delta E_v[\rho] = \int \frac{\delta E_v}{\delta \rho_s^{\text{ref}}(\mathbf{r}, \mathbf{r}')} \frac{\delta \rho_s^{\text{ref}}(\mathbf{r}, \mathbf{r}')}{\delta \rho(\mathbf{r}'')} \delta \rho(\mathbf{r}'') \, d\mathbf{r} \, d\mathbf{r}' \, d\mathbf{r}'' \quad (2.21)$$

$$= \int H_s^{\text{ref}}(\mathbf{r}', \mathbf{r}) \delta \rho_s^{\text{ref}}(\mathbf{r}, \mathbf{r}') \, d\mathbf{r} \, d\mathbf{r}', \quad (2.22)$$

where the auxiliary Hamiltonian is defined as

$$H_s^{\text{ref}}(\mathbf{r}', \mathbf{r}) = \frac{\delta E_v[\rho_s^{\text{ref}}]}{\delta \rho_s^{\text{ref}}(\mathbf{r}, \mathbf{r}')}. \quad (2.23)$$

Therefore, optimization with respect to ρ_s^{ref} is necessary for the minimum required in Eq. (2.3), namely $\delta E_v[\rho] = 0$. The energy gradient, Eq. (2.23), allows the optimization in terms of ρ_s^{ref} . Another approach is to carry out the optimization in terms of orbitals $\{\phi_i(\mathbf{r})\}$, which constitutes the auxiliary non-interacting density matrix ρ_s^{ref} :

$$\rho_s^{\text{ref}} = \sum_i |\phi_i\rangle \langle \phi_i|, \quad (2.24)$$

using the energy gradient

$$\frac{\delta E_v[\rho]}{\delta \phi_i^*(\mathbf{r})} = \int d\mathbf{r}' H_s^{\text{ref}}(\mathbf{r}, \mathbf{r}') \phi_i(\mathbf{r}'). \quad (2.25)$$

There are different ways to obtain the SCF solution. We here develop the generalized optimized effective potential (GOEP) method[34] to indirectly vary ρ_s^{ref} , by optimizing a non-local effective potential v^{GOEP} until

$$\frac{\delta E_v}{\delta (v^{\text{GOEP}})_{ai}} = 0, \quad (2.26)$$

where $(v^{\text{GOEP}})_{ai}$ stands for the occupied-virtual block of the non-local potential.

Within our approach, excited states are treated on the same footing as the ground state. To calculate the energy of the excited state, we minimize the excited-state energy functional, Eq. (2.2) with respect to ρ_s^{ref} . The excited-state information, namely the stationary nature, not the energy minimum, is built in $\Delta E_n[\rho_s^{\text{ref}}]$, the excitation energy part of the functional. The minimization with respect to ρ_s^{ref} is justified, because even in the calculation of excitation energies, the auxiliary system remains in a ground state of the effective Hamiltonian H_s^{ref} . The minimization with respect to ρ_s^{ref} is a part of the search for the optimal auxiliary system for the stationary state (the excited states), the other part being the construction of the excitation energy functional $\Delta E_n[\rho_s^{\text{ref}}]$.

We now consider two specific functionals. Both the pp-RPA and the SF-TDDFT have been carried out as post-DFT approaches. Here we need to reformulate them as a functional of the non-interacting auxiliary system density matrix ρ_s^{ref} . Consider an N -electron physical system, within the pp-RPA[26], $E_v^{\text{ref}}[\rho_s^{\text{ref}}]$ is the energy of the corresponding $(N - 2)$ -electron auxiliary system described with DFA. The two-electron addition excitation energy from the pp-RPA reads

$$\Delta E_n^{+2e} = \begin{bmatrix} (X^{+2e,n})^\dagger & (Y^{+2e,n})^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\text{T} & \mathbf{C} \end{bmatrix} \begin{bmatrix} X^{+2e,n} \\ Y^{+2e,n} \end{bmatrix}, \quad (2.27)$$

where $X^{+2e,n}$ and $Y^{+2e,n}$ are eigenvectors of the pp-RPA matrix, which correspond to the n -th two-electron addition state. The \mathbf{A} and \mathbf{C} matrices are rewritten into non-canonical forms[33]:

$$\mathbf{A}_{ab,cd} = \langle \phi_a | \hat{F} | \phi_c \rangle \langle \phi_b | \phi_d \rangle + \langle \phi_b | \hat{F} | \phi_d \rangle \langle \phi_a | \phi_c \rangle + \frac{1}{2} \langle ab || cd \rangle, \quad (2.28)$$

$$\mathbf{C}_{ij,kl} = -\langle\phi_i|\hat{F}|\phi_k\rangle\langle\phi_j|\phi_l\rangle - \langle\phi_j|\hat{F}|\phi_l\rangle\langle\phi_i|\phi_k\rangle + \frac{1}{2}\langle ij||kl\rangle, \quad (2.29)$$

where \hat{F} is the non-interacting (G)KS Hamiltonian for the auxiliary system:

$$\hat{F} = \frac{\delta E_v^{\text{ref}}[\rho_s^{\text{ref}}]}{\delta \rho_s^{\text{ref}}}. \quad (2.30)$$

The \mathbf{B} matrix is

$$\mathbf{B}_{ab,ij} = \frac{1}{2}\langle ab||ij\rangle. \quad (2.31)$$

Similarly, with the SF-TDDFT, $E_v^{\text{ref}}[\rho_s^{\text{ref}}]$ is the energy of the corresponding high-spin auxiliary system described with DFA. The spin-flip excitation energy from the collinear SF-TDDFT with the Tamm-Dancoff approximation (TDA) is

$$\Delta E_n^{\text{SF}} = (X^n)^\dagger \mathbf{A} X^n, \quad (2.32)$$

where X^n is the eigenvector corresponding to the n -th spin-flipped state. The non-canonical form of the \mathbf{A} matrix here is (following what has been done in Refs 35, 36, 37 for the TDDFT/TDA.)

$$\mathbf{A}_{\bar{a}i,\bar{b}j} = \langle\phi_{\bar{a}}|\hat{F}|\phi_{\bar{b}}\rangle\langle\phi_i|\phi_j\rangle - \langle\phi_i|\hat{F}|\phi_j\rangle\langle\phi_{\bar{a}}|\phi_{\bar{b}}\rangle - c_{\text{HF}}\langle i\bar{b}|\bar{j}\bar{a}\rangle, \quad (2.33)$$

where c_{HF} is the percentage of the exact exchange in the functional used. The bars on top of the spatial orbital indices denotes β spins. It is easy to see that ΔE_n^{+2e} and ΔE_n^{SF} are rotational invariant, as long as the orbital rotation does not mix the occupied and virtual spaces. Thus given a density matrix ρ_s^{ref} , the total energy is uniquely determined. The multireference DFT functionals generated are denoted pp-SCF and SF-SCF, while the original methods are abbreviated as pp and SF. We add “@DFA” after the name to indicate which DFA is used as the reference functional. For

example, “pp@HF” stands for the regular pp-RPA with the Hartree-Fock functional. For the sake of simplicity, we omit this suffix when using the Hartree-Fock reference functional.

2.2.3 Analytic gradients for the pp-RPA and the SF-TDA

The non-zero gradients for the pp-RPA two-electron addition energy part are

$$\begin{aligned} \frac{\delta \Delta E_n^{+2e}}{\delta (v^{\text{GOEP}})_{ia}^*} = \frac{1}{\epsilon_a - \epsilon_i} & \left\{ 2 \sum_{a'b'} (X_{aa'}^{+2e,n})^* F_{ib'} X_{b'a'}^{+2e,n} + \sum_{a'b'c'} (X_{aa'}^{+2e,n})^* \langle ia' || b'c' \rangle X_{b'c'}^{+2e,n} \right. \\ & - 2 \sum_{a'b'c'} (X_{a'b'}^{+2e,n})^* K_{a'c',ai}^{\text{Hxc}} X_{c'b'}^{+2e,n} + \sum_{a'i'j'} (X_{aa'}^{+2e,n})^* \langle ia' || i'j' \rangle Y_{i'j'}^{+2e,n} \\ & - \sum_{a'b'i'} (X_{a'b'}^{+2e,n})^* \langle a'b' || ai' \rangle Y_{ii'}^{+2e,n} + 2 \sum_{i'j'k'} (Y_{i'j'}^{+2e,n})^* F_{i'a} Y_{ij'}^{+2e,n} \\ & \left. + 2 \sum_{i'j'k'} (Y_{i'j'}^{+2e,n})^* K_{i'k',ai}^{\text{Hxc}} Y_{k'j'}^{+2e,n} - \sum_{i'j'k'} (Y_{i'j'}^{+2e,n})^* \langle i'j' || ak' \rangle Y_{ik'}^{+2e,n} \right\} \end{aligned} \quad (2.34)$$

$$= \left(\frac{\delta \Delta E_n^{+2e}}{\delta (v^{\text{GOEP}})_{ia}^*} \right)^*, \quad (2.35)$$

where K^{Hxc} is the Hartree-exchange-correlation kernel.

The non-zero gradients for the spin-flip excitation energy part are (if the orbitals are real and X^n is real)

$$\begin{aligned} \frac{\delta \Delta E_n^{\text{SF}}}{\delta (v^{\text{GOEP}})_{ai}} = \frac{1}{\epsilon_a - \epsilon_i} \sum_{\bar{c}\bar{k}} X_{\bar{c}\bar{k}}^n & \left\{ F_{ka} X_{\bar{c}i}^n + \sum_l K_{kl,ai}^{\text{Hxc}} X_{\bar{c}l}^n \right. \\ & \left. + \sum_{\bar{d}} c_{\text{HF}}(ka|\bar{c}\bar{d}) X_{\bar{d}i}^n - K_{\bar{c}\bar{d},ai}^{\text{Hxc}} X_{\bar{d}k}^n \right\}, \end{aligned} \quad (2.36)$$

$$\begin{aligned} \frac{\delta \Delta E_n}{\delta(v^{\text{GOEP}})_{\bar{a}\bar{i}}} &= \frac{1}{\epsilon_{\bar{a}} - \epsilon_{\bar{i}}} \sum_{\bar{c}k} X_{\bar{c}k}^n \left\{ F_{\bar{i}\bar{c}} X_{\bar{a}k}^n - \sum_{\bar{d}} K_{\bar{c}\bar{d},\bar{a}\bar{i}}^{\text{Hxc}} X_{\bar{d}k}^n \right. \\ &\quad \left. + \sum_l K_{kl,\bar{a}\bar{i}}^{\text{Hxc}} X_{\bar{c}l}^n - c_{\text{HF}}(kl|\bar{c}\bar{i}) X_{\bar{a}l}^n \right\}, \end{aligned} \quad (2.37)$$

Because the restricted open-shell reference is used, the α and β spin gradients are added together to keep the orbitals restricted.

The evaluation of each gradient requires formally $\mathcal{O}(N^3)$ operations, and there are $\mathcal{O}(N^2)$ gradients. Therefore, the overall formal scaling of the optimization is $\mathcal{O}(N^5)$.

2.3 Results

To illustrate the ability of this method to describe systems with multireference characters, we tested the methods with single bond dissociation and double bond rotation. Excitation energies in atoms and small molecules are computed as well. In this work, pp and pp-SCF calculations are performed with restricted singlet references with two electrons removed, followed by spin-adapted calculations. For the SF and SF-SCF, restricted open-shell high-spin triplet references are used. The target states are either singlet or $M_z = 0$ triplet. Since the SF-TDA is spin-incomplete, there will be spin contamination in the target states. The spin contamination observed in SF-SCF calculations is smaller than or comparable to that produced by the SF. Generally, the SCF density matrix for the reference DFA calculation is suitable to be used as the initial guess for ρ_s^{ref} . However, there are cases that this initial guess leads the optimization to be trapped in a local minimum. In these cases, density matrices from converged calculations of similar structures (bond stretch or shrink) may be utilized. The auxiliary system energy is computed with the Hartree-Fock functional unless the DFA functional is specified otherwise. Our methods have been implemented in the

QM4D[38] package. Coupled-cluster singles and doubles (CCSD)[39] calculations for single bond molecules and ethylene are performed with Gaussian 09[40].

The proper description of single bond breaking is challenging for single reference methods. It is well-known that restricted (G)KS DFT/Hartree-Fock produces a wrong dissociation limit. Although unrestricted (G)KS DFT/Hartree-Fock is reasonable in the energy of dissociated atoms, it has serious spin contamination problems and incorrect spin densities when the nuclear separation is large. Multireference methods are capable of predicting the correct dissociation limit. However, the conventional SF-TDA/SF-CIS usually underestimates the binding energy[41, 42, 43, 44, 45]. The pp-RPA could suffer from the problem that orbitals of the $(N - 2)$ -electron system are too contracted[33], leading to underestimated equilibrium bond lengths for the ground state. We expect the optimization of density in our methods to improve the description. Dissociation curves of H_2 , LiH, BH and HF molecules are shown in Figure 2.1, with pp, pp-SCF, SF and SF-SCF results. CCSD results are also shown for reference, except in H_2 . In the case of H_2 , the pp-RPA result is exact and the $N - 2$ density is simply zero, thus making the self-consistent calculation unnecessary[26]. Binding energies and equilibrium bond lengths are presented in Tables 2.1 and 2.2. It can be seen that the description for single bond dissociation is significantly improved. Both the dissociation energy and the equilibrium bond length become more accurate after achieving self-consistency. The error in the equilibrium bond length is greatly reduced as the orbitals of the $(N - 2)$ -electron reference relax during the variation. For the LiH molecule, the pp-SCF curve coincides with the CCSD result. The pp-RPA predicts a state crossing around 1.2 Å for the HF molecule, which does not exist in the CCSD calculation. The pp-SCF corrects this behavior, making the curve relatively similar to the CCSD result.

The choice of the linear response theory affects the overall accuracy. Firstly, the

linear response theory selected determines what configurations are considered in the target state. Secondly, even for conventional post-SCF excitation energy calculations, the SF-TDA and the pp-RPA certainly differ. The optimization is expected to improve the quality of the total energy over the parent linear response theory, thus the overall accuracy relies on the method chosen.

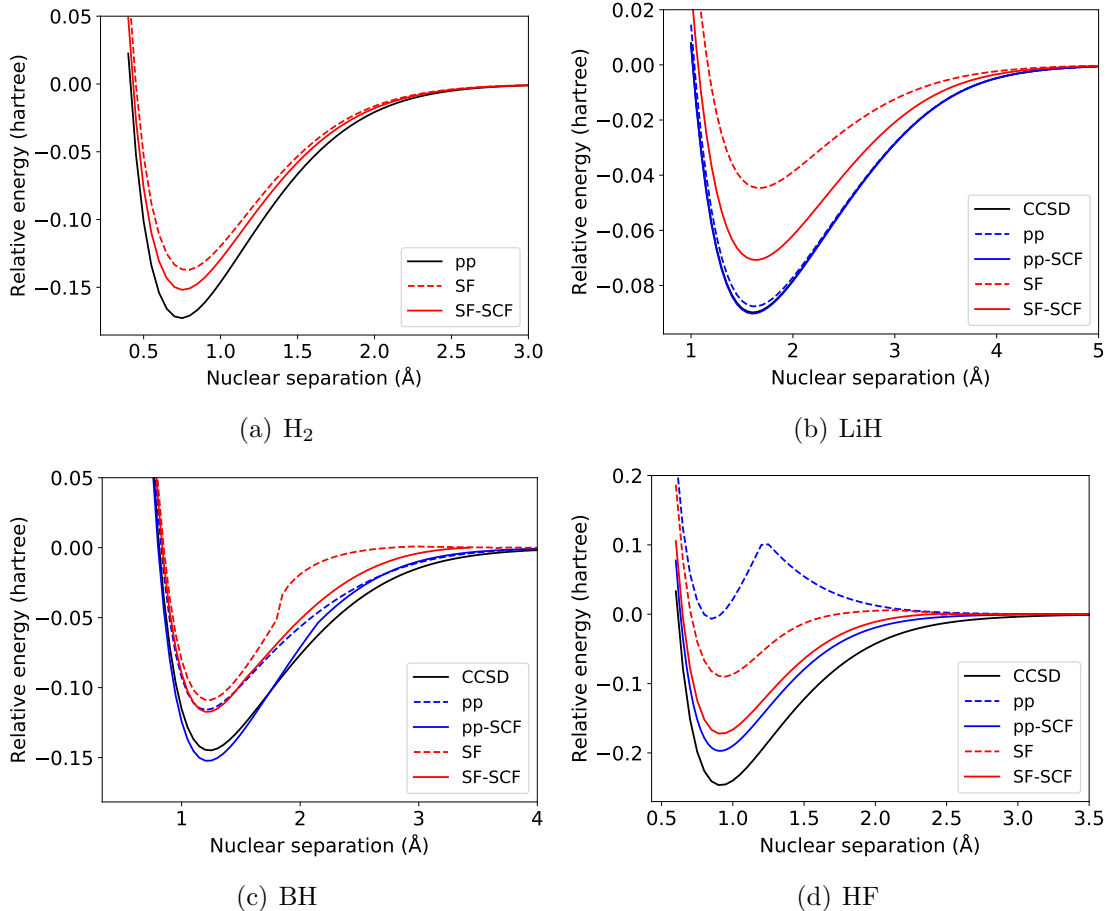


Figure 2.1: Relative energies of H₂, LiH, BH and HF molecules with respect to corresponding dissociated atoms. The cc-pVTZ basis set has been used.

Although the Hartree-Fock functional is mostly used in this work as the reference DFA, subsequent works using functionals which have a good performance can follow straightforwardly. With the Hartree-Fock reference, the SF-TDA reduces to SF-CIS[41, 42], which makes SF-SCF@HF equivalent to MCSCF with spin-flipped

Table 2.1: Binding energies (eV) for H₂, LiH, BH and HF computed with the cc-pVTZ basis set.

Molecule	CCSD	pp	pp-SCF	SF	SF-SCF
H ₂	—	4.70	—	3.73	4.13
LiH	2.44	2.39	2.46	1.22	1.91
BH	3.94	3.15	4.15	2.96	3.20
HF	6.71	0.18	5.36	2.45	4.69
MAD ^a	—	2.46	0.53	1.86	0.96

^a Mean absolute deviations from CCSD values. For H₂, the standard value is chosen as the pp-RPA value. There is no error for pp and pp-SCF in this case, so MAD's for pp and pp-SCF are only computed for LiH, BH and HF.

Table 2.2: Equilibrium bond lengths (Å) for H₂, LiH, BH and HF computed with the cc-pVTZ basis set.

Molecule	CCSD	pp	pp-SCF	SF	SF-SCF
H ₂	—	0.745	—	0.779	0.758
LiH	1.609	1.625	1.608	1.673	1.637
BH	1.236	1.209	1.223	1.225	1.219
HF	0.917	0.857	0.911	0.936	0.919
MAD ^a	—	0.034	0.007	0.032	0.015

^a Mean absolute deviations from CCSD values. For H₂, the standard value is chosen as the pp-RPA value. There is no error for pp and pp-SCF in this case, so MAD's for pp and pp-SCF are only computed for LiH, BH and HF.

configurations from the high-spin configuration. However, even with the Hartree-Fock reference, the pp-SCF@HF energy is not equal to an expectation value of a multiconfigurational wave function on the Hamiltonian. This makes pp-SCF different from MCSCF. Figure 2.2 shows an example where a DFT reference (the HFLYP functional, 100% Hartree-Fock exchange, 100% LYP correlation[14]) is used. Compared

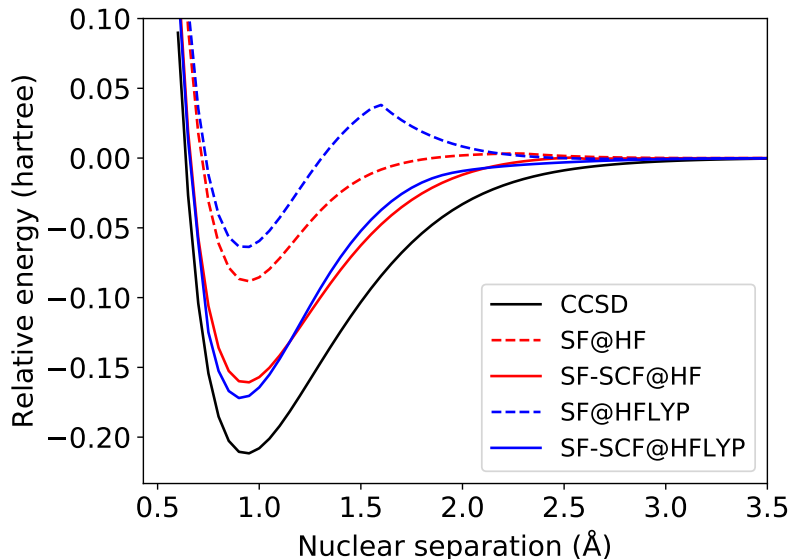


Figure 2.2: Relative energy of HF molecule with respect to dissociated atoms. The 6-31G* basis set has been used.

to SF-SCF@HF, SF-SCF@HFLYP gives slightly better binding energy, which may be attributed to the fact that DFT references provide dynamic correlation, in addition to the static correlation contribution which is already included by the multireference calculation.

The double bond rotation, for example, the ethylene torsion, is another example that static correlation contribution is significant. In this case, even the CCSD produces an unphysical cusp, because the Hartree-Fock calculation is already qualitatively incorrect. Unlike the N -electron singlet ethylene, the $(N - 2)$ -electron singlet for the pp-RPA and the high spin triplet for spin-flip methods are single determinan-

tal. This fact makes the pp-RPA and the SF-TDDFT suitable for this problem. As shown in Figure 2.3, they will not give an unphysical cusp, but the barrier height is not optimal. Self-consistency corrects the barrier heights to be closer to the MR-ccCA value[46].

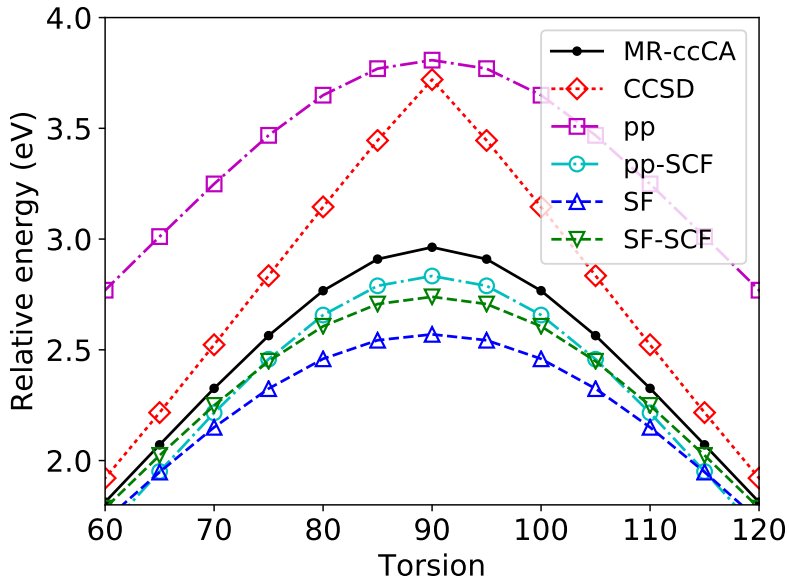


Figure 2.3: Relative energy of twisted ethylene with respect to flat ethylene. The cc-pVTZ basis set has been used. Structures are adapted from Ref. 46. The carbon-carbon distance varies when the torsion angle θ changes, $r_{CC} = (1.339 + 0.12\theta/90)$ Å. Other parameters are fixed, $r_{CH} = 1.0856$ Å, $\angle HCH = 121.2^\circ$.

These results demonstrate that the pp-SCF and SF-SCF are good ground state multireference energy functionals. Since their parent methods are mostly used to calculate excitation energies, it is also worthwhile to investigate their performance on excited states. Atomic excitation energies and molecular vertical excitation energies are presented in Tables 2.3 and 2.4. For atomic excitation energies, self-consistency seems not to change the results much, especially for the SF-SCF. Larger modifications occur when the pp-SCF calculates excitation energies for the oxygen atom. For molecular excitations, The pp-SCF outperforms the pp-RPA on singlet-triplet gaps, but singlet-singlet gaps seem to get overestimated. The SF-SCF generally performs

better than the non-self-consistent SF.

Table 2.3: Atomic excitation energies (eV). The cc-pVTZ basis set has been used in the calculation for Ca, while the aug-cc-pVTZ basis set has been used for other calculations. Standard values are: Be: experimental values from Ref. 47. Mg: experimental values from Ref. 48. Ca: computed values from CR-EOM-CCSD(T)[49]. O and S: experimental values from Ref. 50.

Term	Standard	pp	pp-SCF	SF	SF-SCF
Be					
³ P	2.73	2.74	2.74	2.06	2.11
¹ P	5.28	5.34	5.34	5.60	5.25
Mg					
³ P	2.71	2.58	2.59	2.07	2.14
¹ P	4.35	4.27	4.28	4.49	4.17
Ca					
³ P	1.79	1.66	1.68	1.26	1.32
O					
¹ D	1.97	1.75	2.02	2.09	2.04
¹ S	4.19	2.98	3.46	4.61	4.54
S					
¹ D	1.15	1.20	1.30	1.21	1.15
¹ S	2.75	1.95	2.09	3.07	3.02
MAD ^a	—	0.30	0.22	0.36	0.28

^a Mean absolute deviations from standard values.

Table 2.4: Molecular vertical excitation energies (eV). The aug-cc-pVDZ basis set has been used for the calculations of ethylene excitation energies. For other molecules, the aug-cc-pVTZ basis set has been used. Standard values: BH and CH⁺: computed values using CR-EOM-CCSD(T)[49]. CO and N₂: experimental values from Ref. 51. Ethylene: computed values from Ref. 52.

Term	Standard	pp	pp-SCF	SF	SF-SCF
BH					
³ Π	1.27	1.66	1.61	0.77	1.03
¹ Π	2.85	3.18	3.52	3.17	2.95
CH ⁺					
³ Π	1.15	1.72	1.59	0.54	0.87
¹ Π	3.07	3.60	3.91	3.21	3.09
CO					
³ Π	6.32	5.63	6.47	5.42	6.17
¹ Π	8.51	7.84	9.29	8.98	9.23
N ₂					
³ Σ _u ⁺	7.75	8.32	9.03	7.96	7.51
¹ Π _g	9.31	9.81	10.52	9.72	9.51
Ethylene					
³ B _{1u}	4.5	3.94	4.48	—	—
¹ B _{1u}	7.8	6.28	8.69	—	—
MAD ^a	—	0.63	0.66	0.44	0.24

^a Mean absolute deviations from standard values.

Chapter 3

Machine-Learn a Density Functional for Optimal Accuracy

3.1 Introduction

Among many electronic structure methods, the density functional theory (DFT) [53, 54] has won the widest application because of its favorable accuracy and reasonable cost. Although the DFT is formally exact, its success is attributed to the Kohn-Sham scheme [54] and approximations to the exchange-correlation energy (E_{xc}).

Many semi-local density functionals approximations (DFAs) have been developed during recent decades [55, 56, 57, 58, 59, 60]. Great success has been achieved by semi-local approximations, while challenges still present. Semi-local DFAs are known to have large delocalization errors [61, 62]. Besides, long-range correlation effects are difficult to be captured within a local form. This would lead to, for example, the inability on describing the van der Waals interaction [63]. Even in applications where above issues are considered insignificant, semi-local DFAs may still produce inaccurate results. Systematic improvements of (meta-)GGAs with data indicate that, the local form might limit the accuracy that semi-local functionals can achieve [64]. Non-local components in the functional are necessary to overcome these difficulties.

Non-local functionals have already been developed and widely used for many years. One way to introduce non-locality into the functional is to mix Hartree-Fock non-local exchange into a semi-local functional, making the resulting approximation

a hybrid functional [65, 66]. Hybrid functionals partially reduce the delocalization error [62], and are generally more accurate than semi-local functionals. In recent years, double-hybrid functionals have attracted much research interest [67, 68, 69, 70] In addition to the non-local exchange, double-hybrid functionals use non-local correlation energies, which are usually derived from many-body perturbation theories, such as the second-order Møller-Plesset perturbation theory (MP2) and the random phase approximation (RPA). Double-hybrid functionals outperform approximations in other categories, thus making them in the highest rung of the Jacob’s ladder.

Though hybrid and double-hybrid functionals have appealing accuracy, the relatively high computational cost may prohibit their practical applications, especially when used to calculate large molecules and periodic systems. On the other hand, even the best-performing hybrid or double-hybrid functionals are still not guaranteed to reach the chemical accuracy [71]. It remains a challenge on how to further improve hybrid and double-hybrid functionals, as the non-local parts do not have much flexibility to adjust. Is there a way to design a non-local DFA such that orbitals are not required and the performance is possible to be systematically improved?

The main challenge is that, the non-local exchange-correlation kernel is much harder to approximate than a local exchange-correlation energy density. Only in very rare instances can the non-local kernel be explicitly derived. The VV10 non-local correlation energy [72] is an example, while it tackles the non-covalent interaction only. To the best of our knowledge, for a general purpose non-local density functional, there is no existing mathematical expression to build upon. Has creating a non-local density functional become an impossible mission? Fortunately, the many-body expansion, along with machine-learning techniques, can help to create such a functional. The many-body expansion simplifies the development of a non-local functional to finite-range non-local sub-problems, which is a requirement for directly

machine-learning the non-local functional. In following sections, details about the many-body expansion and the machine-learning techniques will be presented.

3.2 Many-body Expansion in the DFT

The many-body expansion is commonly used in computational chemistry to reduce the computational cost. Its success has been witnessed in fragment based methods, [73, 74, 75] and the method of increments. [76, 77] In terms of the total electron density, a similar definition for the many-body expansion can also be made: a part of the total energy is expressed as a summation of one-body energies, two-body energies and terms that consist of more bodies,

$$\Delta E[\rho] = \sum_i \Delta E_1[\rho_i] + \sum_{i<j} \Delta E_2[\rho_i, \rho_j] + \sum_{i<j<k} \Delta E_3[\rho_i, \rho_j, \rho_k] + \dots, \quad (3.1)$$

where ρ is the total electron density, and ρ_i is the density for one body. The choice of the one-body density ρ_i is crucial for the quality of the many-body expansion. One-body densities should be local, such that the many-body expansion can converge. It is natural to define the body as the local density around the nuclei. In this work, the Hirshfeld partition scheme [78] is used to generate the local densities:

$$\rho_i(\mathbf{r}) = \frac{\rho_i^0(|\mathbf{r} - \mathbf{R}_i|)}{\sum_a \rho_a^0(|\mathbf{r} - \mathbf{R}_a|)} \rho(\mathbf{r}), \quad (3.2)$$

where ρ_a^0 is the spherical electron density of free atom a . In Fig. 3.1 an example of the partition of the electron density is shown.

It is also important to determine which part of the total energy should be approximated by the many-body expansion. The total energy, or even the much smaller exchange-correlation energy, may require many terms in the many-body expansion

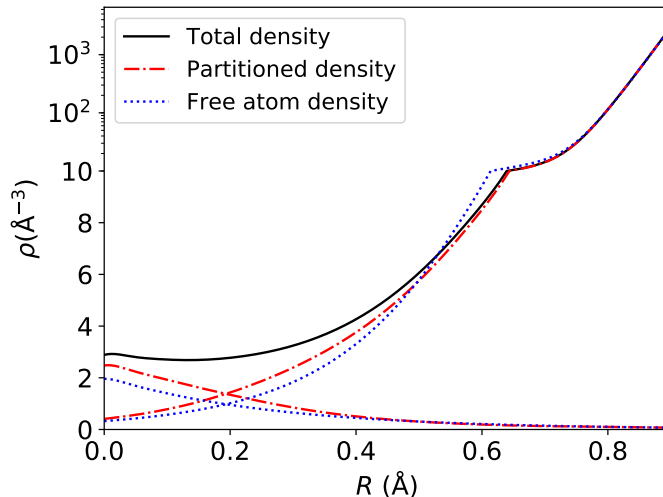


Figure 3.1: The electron density along the chemical bond of the HF molecule. The nuclear separation is 0.9 \AA .

to achieve convergence. As shown in Fig. 3.2, the one-body expansion approximation fails to recover the BLYP exchange-correlation energy. A smaller quantity, the difference between total energies produced by different DFAs, is usually safe to truncate at one-body terms. Therefore, we express the DFT total energy as

$$E[\rho] \approx E^{\text{base}}[\rho] + \sum_i \Delta E_1[\rho_i], \quad (3.3)$$

where E is the exact energy, and E^{base} is the ground level we start from. As pointed out before, E^{base} should already contain an approximate E_{xc} (usually a semi-local functional), otherwise terminating at one-body terms will not suffice. Even if in Fig. 3.2 we have shown that, with the aid of an one-body correction, E^{BLYP} can be recovered from a LDA calculation, this observation, however, does not necessarily apply to the exact energy. Long-range correlation effects, such as the van der Waals interaction, may require two and more body terms, if the base functional does not fully capture the long-range correlation. In this work, we leave the two-body van

der Waals interaction to the DFT-D3 model, [79, 80] and other many-body interactions (many-body van der Waals interactions, [81] fractional charge errors and static correlation) are not considered.

The usage of one-body correction is based on the observation that the correlation effects are mostly “local”. However, it does not imply that the resulting functional is local. In fact, $\Delta E_1[\rho_i]$ is relatively “local” as comparing to the entire chemical system. Its functional form, however, can be completely non-local in terms of the local density ρ_i . Hence, each one-body term $\Delta E_1[\rho_i]$ may be called a finite-range non-local functional. The finite-range character comes from the fact that ρ_i resides only in a certain range. Beyond that range, ΔE_1 can hardly play a role. This greatly simplifies the procedure of approximating ΔE_1 . By expressing the approximate functional as a base semi-local functional plus one-body corrections, although the resulting functional is not completely non-local, it already goes beyond traditional semi-local functionals. The combination of the locality in the atom scale, and the non-locality in terms of the functional form, enables the development of a density functional that captures most of the short-range correlation effects, without needing virtual orbitals.

It is easy to see that, since each sub-problem is small and independent of the system size, the cost for evaluating the one-body correction is linearly scaled. Therefore, accurate (and probably expensive at the same time) methods that take the density as the input, such as the direct optimization method, [82] can be used without making this functional unaffordable. However, reducing the computational cost is not the only reason to use the many-body expansion. After reducing to atom-centered sub-problems, the energy correction automatically becomes translational invariant and size-extensive. This removes the restrictions that one may encounter when designing a global non-local functional, thus making it possible to machine-learn the exchange-correlation energy functional (by learning the one-body energy correction). This is

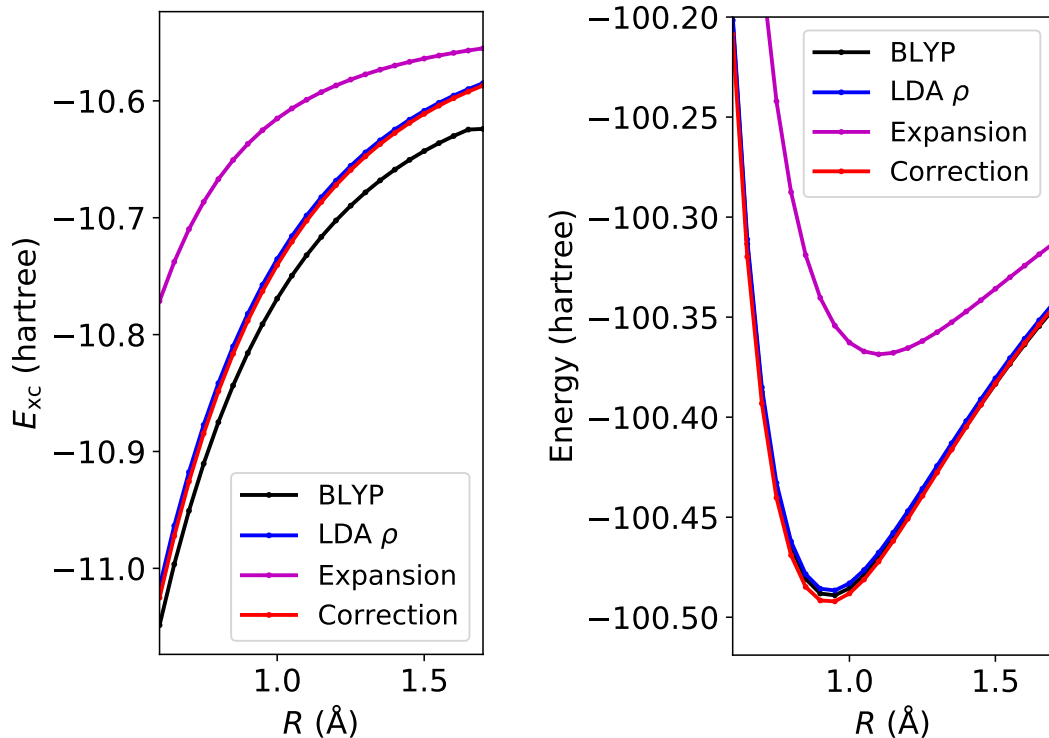


Figure 3.2: (Left) Exchange-correlation energy in HF for nuclear separations from 0.6 Å to 1.7 Å. The “BLYP” curve means E_{xc}^{BLYP} evaluated with BLYP self-consistent density (ρ^{BLYP}). The “LDA ρ ” curve is E_{xc}^{BLYP} evaluated with LDA self-consistent density (ρ^{LDA}). The curve labeled as “Expansion” means we directly decompose E_{xc} using partitioned densities: $E_{xc} = E_{xc}^{\text{BLYP}}[\rho_{\text{H}}^{\text{LDA}}] + E_{xc}^{\text{BLYP}}[\rho_{\text{F}}^{\text{LDA}}]$. The “Correction” curve is computed by using $E_{xc} = E_{xc}^{\text{LDA}}[\rho^{\text{LDA}}] + (E_{xc}^{\text{BLYP}} - E_{xc}^{\text{LDA}})[\rho_{\text{H}}^{\text{LDA}}] + (E_{xc}^{\text{BLYP}} - E_{xc}^{\text{LDA}})[\rho_{\text{F}}^{\text{LDA}}]$. (Right) Total energy of HF molecule computed in the ways as discussed before. It is noteworthy that even there is a difference in the exchange-correlation energy when using different densities, the total energy still agrees very well. This can be attributed to the variational property of the total energy.

similar to machine-learning the molecular energy by adding up atomic contributions in Ref. [83].

3.3 Machine-Learn the Finite-Range Non-Local Functional

Even if the sub-problems seems to be simple, an explicit functional form is still impractical to be derived for $\Delta E_1[\rho_i]$. Machine-learning happens to be an excellent tool to solve for this finite-range non-local functional numerically. Machine-learning has already been applied in quantum chemistry, [84, 85], and particularly, in the field of DFT. [86, 87] In this work, the sub-problem is suitable to be solved by machine-learning, because of the finite-range character of the local density ρ_i . ρ_i is an atom-centered density that is partitioned from the molecular total density, and in most cases it is not too different from the electron density of free atom i . Therefore, by feeding high quality data for various element types to the machine-learning model, it is possible that this one-body energy correction can be predicted for any given input local density (in the interpolation range), with a quite low computational cost. Then the correction to the total energy of the chemical system is a simple summation of atomic contributions, which is ensured by the many-body expansion.

Designing an approximate functional by machine-learning is very distinct from what have been done for conventional functionals. Currently available functionals all strive for as few parameters as possible, and are often fitted to exact constraints and limited number of data. In this way, the functional developed hopefully can be applicable in most systems. By contrast, machine-learning methods usually have a huge amount of parameters, rely heavily on the training data, and are expected

to produce good interpolation quality. It looks like machine-learning is not relevant to functional development, however, it is the black box nature of machine-learning that makes the development of the non-local density functional possible. As long as enough training data are provided, there is no prior knowledge required on the form of the function fitted. The use of data in the functional development has already been well established in the development of functionals like Minnesota functionals [88, 89] and ω -B97 series [90, 91]. Machine-learning is simply another tool to utilize the data.

To express the finite-range non-local functional as a machine-learning function, the local density needs to be pre-processed to provide finite number of inputs. The implementation details are provided in the next section.

3.4 Methods

A conventional GGA needs to be provided as the base functional E_{xc}^{base} . It is used not only to provide the base energy, but also to produce the density. Currently the implementation does not modify the density when evaluating the machine-learning functional, thus this is a post-SCF method. The goal of the energy correction is to reproduce accurate energies that are generated by high-level theories:

$$E^{\text{accurate}}[\rho(\mathbf{r})] = E^{\text{base}}[\rho(\mathbf{r})] + \Delta E^{\text{ML}}[\rho(\mathbf{r})], \quad (3.4)$$

where the energy produced by the base method is

$$E^{\text{base}}[\rho] = \sum_{A<B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} - \int \rho(\mathbf{r}) \sum_A \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + T_s[\rho(\mathbf{r})] + E_{\text{xc}}^{\text{base}}[\rho(\mathbf{r})]. \quad (3.5)$$

The machine-learning energy correction is a correction to the exchange-correlation energy, and it is approximated as a summation of atomic contributions:

$$\Delta E^{\text{ML}} = \Delta E_{\text{xc}}^{\text{ML}}[\rho(\mathbf{r})] \approx \sum_i f[\rho_i(\mathbf{r})], \quad (3.6)$$

where f is the actual machine-learning functional that produces a single atom contribution, and ρ_i is the partitioned density.

Many machine-learning models are available and each serves for its particular purpose. In this work, we use the feed-forward artificial neural network (multilayer perceptrons) as an example, and the theory developed here should apply in most other machine-learning models. Because neural networks do not have any built-in symmetry, special care needs to be taken when designing the input. The many-body expansion already takes care of the translational invariance, hence the rotational invariance should be guaranteed by the transformed input. One common choice of making rotational invariant features is using the power spectrum $\{p_{nl}\}$ [92]:

$$p_{nl} = \sum_{m=-l}^l c_{nlm}^* c_{nlm}, \quad (3.7)$$

where c_{nlm} is the coefficient that the density projects on a basis function, which is

centered at atom A :

$$c_{nlm} = \langle \rho_A(\mathbf{r}) | g_n(|\mathbf{r} - \mathbf{R}_A|) Y_l^m(\widehat{\mathbf{r} - \mathbf{R}_A}) \rangle. \quad (3.8)$$

Y_l^m 's are complex spherical harmonics, and g_n is a Gaussian function:

$$g_n(r) = \left(\frac{2\sigma_n}{\pi} \right)^{\frac{1}{4}} \exp[-\sigma_n(r - \mu_n)^2], \quad (3.9)$$

where $\{\mu_n\}$ is the set of nodes that the radial functions are sampled on, and σ_n is determined in the following way:

$$\sigma_n = \begin{cases} \frac{c^2}{(\mu_1 - \mu_0)^2} & n = 0, \\ \frac{4c^2}{(\mu_{n+1} - \mu_{n-1})^2} & 0 < n < n_{\max} - 1, \\ \frac{c^2}{(\mu_n - \mu_{n-1})^2} & n = n_{\max} - 1. \end{cases} \quad (3.10)$$

In this work, $\{\mu_n\}$ is chosen to be the set that includes zero and roots of the Gauss-Laguerre polynomial (linearly scaled to make the largest root become μ_{\max}), and c is an adjustable parameter. Fig. 3.3 shows a typical radial function setup used in this work. Since the creation of the power spectrum for each atom is limited to a finite region around the atom, this further imposes a limit on the range that the non-local functional can explore.

The DFT-D3 energy [79, 80], which is used to partially recover the ignored two-body terms in the many-body expansion, is added to the total energy. The parameters for the D3 model are to be re-optimized during the training process of the neural network.

To handle spin-unrestricted calculations properly, we need to keep the symmetry that $E[\rho^\uparrow, \rho^\downarrow] = E[\rho^\downarrow, \rho^\uparrow]$ (when no external magnetic field is applied). The goal can

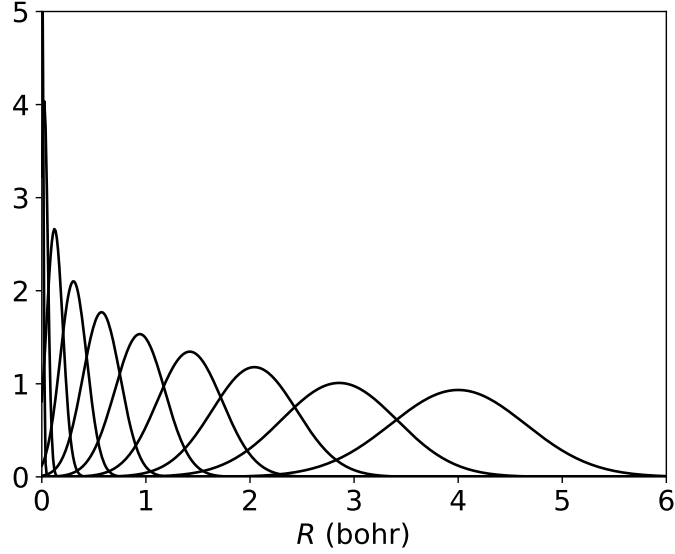


Figure 3.3: An example plot of the radial functions used when generating the power spectrum. In this case, 10 radial functions are generated, with $\mu_{\max} = 4$ bohr and $c = 1.25$.

be easily achieved by using the summation of spin up and down power spectrum as the input for the neural network:

$$\Delta E_{\text{xc}}^{\text{ML}}[\rho^\uparrow, \rho^\downarrow] = \Delta E_{\text{xc}}^{\text{ML}}(\{p_{nl}^\uparrow\}, \{p_{nl}^\downarrow\}) = \Delta E_{\text{xc}}^{\text{ML}}(\{p_{nl}^\uparrow + p_{nl}^\downarrow\}). \quad (3.11)$$

Notice that this is not the same as taking the total density $\rho^\uparrow + \rho^\downarrow$ as the input.

3.4.1 Details about the neural network

The neural network model used in this work is fixed to have 2 hidden layers, thus this model is controlled by 3 weight matrices (\mathbf{W}^0 , \mathbf{W}^1 and \mathbf{W}^2) and 3 bias vectors (\vec{b}^0 , \vec{b}^1 and \vec{b}^2). Hyperbolic tangent is used as the activation function for hidden neurons.

The cost function used to optimize the neural network is defined as:

$$L(\{\mathbf{W}^k\}, \{\vec{b}^k\}) = \frac{1}{N} \sum_i^N \log(\cosh(\hat{y}_i - y_i)) + \beta \sum |w_{ij}^k|, \quad (3.12)$$

where N is the number of entries in the database, \hat{y} is the reference value and y is the value produced by the neural network, given the set of weights $\{\mathbf{W}^k\}$ and biases $\{\vec{b}^k\}$. The second term in the cost function is the L1 regularization, which is used to prevent over-fitting. Since a single neural network function f produces an energy correction for one atom in a molecule, while reference data are usually reaction energies, in most cases y is a linear combination of atomic contributions. For example, for reaction $X + Y \rightarrow Z$, the reaction energy produced by this functional is

$$y = E_Z^{\text{base}} - E_X^{\text{base}} - E_Y^{\text{base}} + \sum_i^{n_Z} f(\{p_{nl}^i\}) - \sum_j^{n_X} f(\{p_{nl}^j\}) - \sum_k^{n_Y} f(\{p_{nl}^k\}). \quad (3.13)$$

3.4.2 Database rescaling

Since reaction energies in the database can differ in the order of magnitude, the reaction energy data need to be scaled first to avoid ignoring small numbers. We follow the idea of weighted total mean absolute deviations (WTMAD) as presented in Ref. [93]. The reaction energy and stoichiometry for each reaction is scaled by

$$\text{scale} = \frac{81.24 \text{ kcal/mol}}{\overline{|\Delta E_i|}}, \quad (3.14)$$

where $\overline{|\Delta E_i|}$ is the mean reaction energy for subset i that the reaction belongs to. 81.24 kcal/mol is the average value of $\overline{|\Delta E_i|}$ for all subsets.

3.5 Computational Details

The accuracy of the machine-learning model is mostly controlled by the quality of the training data. The data we used in this work are from MGCDB84[71] and GMTKN55[93] databases, which are mostly generated by the coupled-cluster method

[94] (in particular, CCSD(T): a full treatment of singles and doubles, with perturbative triples) in the complete basis set (CBS) limit.

Gaussian 16[95] is used to perform base functional calculations (BLYP and PBE), with a pruned (99, 590) grid and tight convergence criteria (root of mean squared change in the density matrix $< 10^{-8}$). A finer grid (500, 974) is adopted in the calculations of 18 atom absolute energies. Broken-symmetry unrestricted is preferred if it lowers the total energy. The def2-QZVPPD basis set[96] is used to reduce possible basis set superposition error (BSSE). Investigations on the impact of different basis sets are beyond the scope of this work. For calculations that involve elements that are heavier than Kr, empirical core potentials (ECPs) are used. In the computation of the power spectrum, only valence electron densities contribute. Core electron densities that are from the ECPs are not used.

To train and test and neural network, we mainly use reaction energy data that have elements in first four periods. In Fig. 3.4 we show how many times each element appears in the molecules that are used in the database. We reserve 10% of the total data as the test set, while the remaining 90% is used for training. The GMTKN55 database has some data involving main group heavy elements, which will also be used to test the functional.

The generation of the power spectrum from densities is done by an utility program, which is available for downloading at our site. It uses the wave-function file generated by Gaussian to reproduce the density, and computes c_{nlm} numerically with standard DFT grids. Spherical electron densities of free atoms, which are needed in the density partition function, are also computed in this work. The terms in the DFT-D3 (with BJ damping) energy are computed with the `dftd3` utility provided on the DFT-D3 official site, then the parameters, s_6 , s_8 , a_1 and a_2 , are adjusted when the neural network parameters are being optimized. There initial values are taken as the optimal

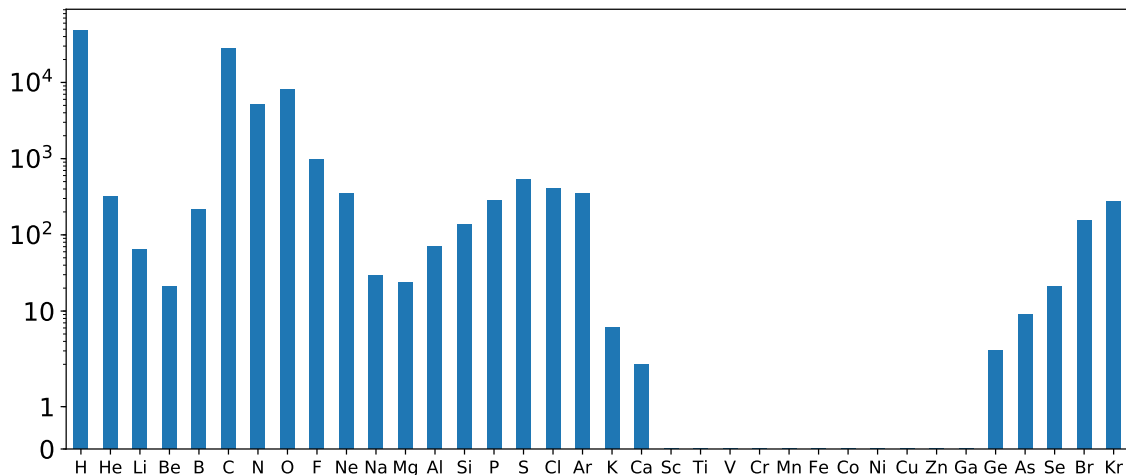


Figure 3.4: The number of times when an element appears in the molecules that are used by the database.

values fitted for the base functional.

A TensorFlow script is used to train the model, which uses the Adam optimizer (with learning rate = 0.001) and early-stopping. The weights of the neural network are initialized with uniform random numbers, and the biases are zero-initialized. 10-fold cross-validation (CV) is used to produce cross-validation errors when tuning the parameters in the input features and hyperparameters in the training process. The hyperparameters (regularization factor β and batch size) are sampled on a simple 3×3 grid, where β can choose from 5×10^{-8} , 7.5×10^{-8} and 1×10^{-7} , and the batch size can be 128, 256 or 512.

The computational cost of our method involves two parts: the cost of creating the power spectrum, and the cost of neural network computations. The cost of evaluating a trained neural network is quite cheap, and there are N neural networks to compute for a molecule with N atoms. The more expensive part is the generation of the power spectrum, which involves numerical integrations. However, it is still much cheaper than the base functional calculation ($\mathcal{O}(N^3)$), because it at most needs

$\mathcal{O}(N^2)$ calculations (there are N atoms, and for each atom, the program needs to loop over the numerical grid of the entire molecule). For large molecules, the cost can be reduced to close to $\mathcal{O}(N)$, since Gaussian radial functions decay very quickly beyond a certain range.

3.6 Results and Discussion

Although a larger neural network usually has a higher capability of learning the unknown function, using an over-sized neural network along with limited number of training data may risk over-fitting. In this work, there are around 5,000 training data points, while the number of parameters in the neural network can easily exceed this quantity. Controlling the number of hidden nodes is one way of limiting the size of the neural network. On the other hand, since we are generating discrete inputs from a continuous function, the neural network does not have a fixed input size. The number of inputs might affect the performance of the neural network more straightforwardly.

The number of the power spectrum reflects how accurately the density is discretized. It is determined by the number of radial functions (n_{\max}) and the angular momentum cut-off in the angular expansion (l_{\max}). Usually the more radial functions are used, and the higher angular momentum cut-off is set, the better quality the discretization will have. However, a high resolution in the power spectrum does not necessarily improve the performance of the neural network. As shown in Fig. 3.5, the contribution from high angular momentums diminishes quickly. This is also reflected in Fig. 3.6(a) that $l_{\max} = 4$ is sufficient for the angular momentum cut-off. n_{\max} has a similar effect on the overall performance, as seen from Fig. 3.6(b). Though larger neural networks and denser power spectrum can produce smaller cross-validation errors, we choose to continue the investigation of the neural network with $n_{\max} = 10$

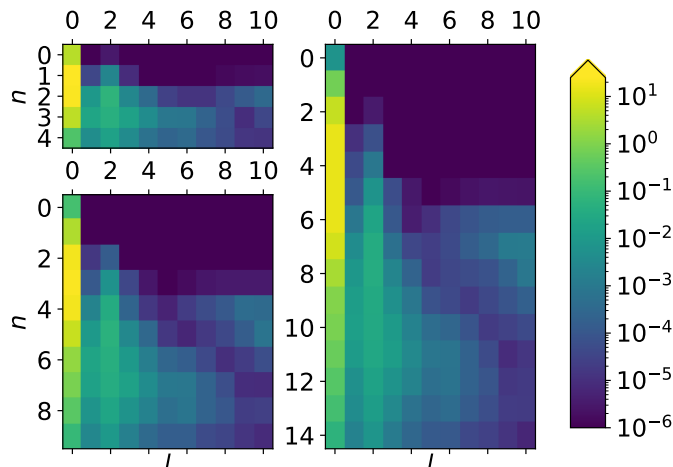


Figure 3.5: Superposition of the power spectrum computed on the entire database, with BLYP densities and parameters $\mu_{\max} = 4$ bohr, $c = 1.25$. 3 figures correspond to $n_{\max} = 5, 10$ and 15 , respectively. l_{\max} is set to 10 in all cases, and the power spectrum with a smaller l_{\max} is simply a sub-graph of the graph with $l_{\max} = 10$.

and 32 neurons in each hidden layer. In this way, we can achieve a relatively good accuracy with a moderate neural network size (50 input nodes and 2,721 parameters in total).

There are two more parameters for input features, μ_{\max} and c , that need to be optimized. They control the form of radial basis functions, and can only be searched by computing the cross-validation errors for a set of values. This process is briefly shown in Fig. 3.7. After the parameters for input features, and hyperparameters in the neural network training process are determined, the neural network is trained against the entire training set. Because initial guesses of neural network weights can affect the training quality, 10 training tasks were run with different starting points. Afterwards, the model (weights and biases of the neural network, and D3 parameters) who has the least training error, is selected from these 10 trained models. To reduce possible random errors produced by neural networks, we then fix the D3 parameters obtained from the best model, and conduct 19 more neural-network-only training

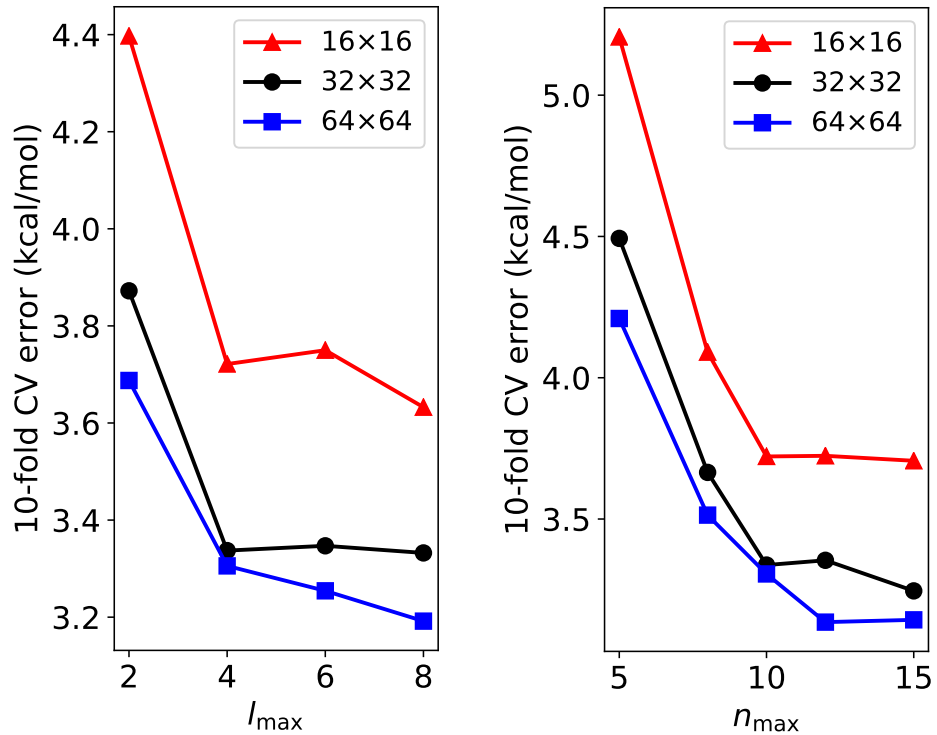


Figure 3.6: (a) Cross-validation errors for different l_{\max} , with BLYP as the base functional. Other parameters: $n_{\max} = 10$, $\mu_{\max} = 4$ bohr and $c = 1.25$. The size of the neural network is reflected in the labels. For example, 16×16 means a neural network with 2 hidden layers, while each hidden layer has 16 hidden nodes. (b) CV errors for varying n_{\max} , while other settings are the same ($l_{\max} = 4$).

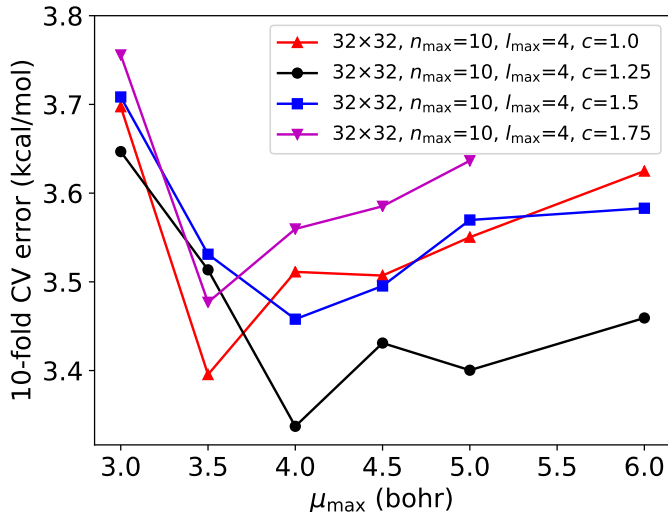


Figure 3.7: Cross-validation errors for a few different setups of μ_{\max} and c , with BLYP base functional.

tasks. Finally, an ensemble of 20 neural networks will be created:

$$\bar{f} = \sum_i w_i f_i, \quad (3.15)$$

where $\{w_i\}$ is the set of ensemble weights, which should fulfill $0 \leq w_i \leq 1$ and $\sum_i w_i = 1$. The ensemble weights are further optimized, such that the lowest possible training error can be obtained. Notice that in the end some of the ensemble weights can be zero, thus the ensemble neural network usually consists of fewer than 20 neural networks.

Table 3.1 and Table 3.2 show the training errors in each individual large data set, with the leading functionals in those sets. But to make a fair comparison, the resulting ensemble neural network and corresponding D3 parameters are tested on the reserved pure test set. Since we are using a mixture of MGCDB84 and GMTKN55, here we separate the tests to distinguish where the tests are from, and compare the neural network functional to different methods for each database. (This is because we want

Table 3.1: Comparing neural network functionals with ω B97M-V for training data belong to MGCDB84. TC: thermochemistry reaction energies. BH: barrier heights. IE: isomerization energies. NC: non-covalent interaction energies. RG: rare gas dimer interaction energies. The unit is kcal/mol.

		MAD	RMSD	WTMAD
TC	BLYP-NN-D3(BJ)	1.57	3.01	1.59
	PBE-NN-D3(BJ)	1.63	3.05	1.70
	ω B97M-V	2.14	2.99	2.78
BH	BLYP-NN-D3(BJ)	1.46	2.02	3.70
	PBE-NN-D3(BJ)	1.07	1.38	3.04
	ω B97M-V	1.06	1.44	3.08
IE	BLYP-NN-D3(BJ)	0.27	0.67	1.94
	PBE-NN-D3(BJ)	0.24	0.60	1.84
	ω B97M-V	0.44	1.01	4.05
NC	BLYP-NN-D3(BJ)	0.11	0.21	2.14
	PBE-NN-D3(BJ)	0.14	0.32	2.29
	ω B97M-V	0.14	0.31	3.06
RG	BLYP-NN-D3(BJ)	0.014	0.028	3.47
	PBE-NN-D3(BJ)	0.023	0.038	5.63
	ω B97M-V	0.048	0.183	11.9

Table 3.2: Comparing neural network functionals with DSD-BLYP-D3(BJ) for training data belong to GMTKN55. Notations are the same as in Table 3.1. The unit is kcal/mol.

		MAD	RMSD	WTMAD
TC	BLYP-NN-D3(BJ)	3.01	4.96	2.53
	PBE-NN-D3(BJ)	3.02	4.97	2.58
	DSD-BLYP-D3(BJ)	2.36	5.49	3.13
BH	BLYP-NN-D3(BJ)	1.04	1.54	4.01
	PBE-NN-D3(BJ)	1.04	1.48	4.01
	DSD-BLYP-D3(BJ)	1.06	1.50	4.21
IE	BLYP-NN-D3(BJ)	2.04	4.64	4.91
	PBE-NN-D3(BJ)	1.68	3.71	4.39
	DSD-BLYP-D3(BJ)	1.94	3.97	7.03
NC	BLYP-NN-D3(BJ)	0.23	0.42	3.02
	PBE-NN-D3(BJ)	0.25	0.52	3.35
	DSD-BLYP-D3(BJ)	0.27	0.44	4.79

Table 3.3: Comparing neural network functionals with ω B97M-V for tests belong to MGCDB84. Notations are the same as in Table 3.1. The unit is kcal/mol.

		MAD	RMSD	WTMAD
TC	BLYP-NN-D3(BJ)	1.90	3.22	1.94
	PBE-NN-D3(BJ)	1.89	3.10	2.23
	ω B97M-V	2.62	3.53	3.49
BH	BLYP-NN-D3(BJ)	1.16	1.40	3.22
	PBE-NN-D3(BJ)	1.14	1.31	3.69
	ω B97M-V	0.76	0.97	2.06
IE	BLYP-NN-D3(BJ)	0.28	0.70	2.17
	PBE-NN-D3(BJ)	0.18	0.32	2.00
	ω B97M-V	0.33	0.49	4.07
NC	BLYP-NN-D3(BJ)	0.16	0.73	2.26
	PBE-NN-D3(BJ)	0.24	1.21	2.76
	ω B97M-V	0.15	0.35	3.04
RG	BLYP-NN-D3(BJ)	0.012	0.019	3.05
	PBE-NN-D3(BJ)	0.019	0.027	4.62
	ω B97M-V	0.013	0.035	3.29

to reuse computed values as provided by these databases, but they did not cover all available functionals, and the best-performing methods ranked by each database are not the same.) For tests in the MGCDB84 database, the comparison with ω B97M-V is presented in Table 3.3. Similarly, Table 3.4 shows the errors on tests that are from the GMTKN55 database, along with the errors produced by DSD-BLYP-D3(BJ). It is intriguing to note that various error indicators (mean absolute deviation (MAD), root of mean squared deviation (RMSD) and WTMAD with scaling factor defined in Eqn. 3.14) can rank the performance of approximate functionals differently. The RMSD is known to overemphasize large absolute errors, which is the reason why neural network functionals have large RMSDs for the TC category in Table 3.4. For example, the ionization potential (IP) of the oxygen atom produced by BLYP-NN-D3(BJ) has an absolute error of 44.7 kcal/mol, while the relative error is actually only 14%. If this term is absent from the test set, the RMSD of the TC category immediately drops to 3.58 kcal/mol (drops by 45%), while the MAD and the WTMAD

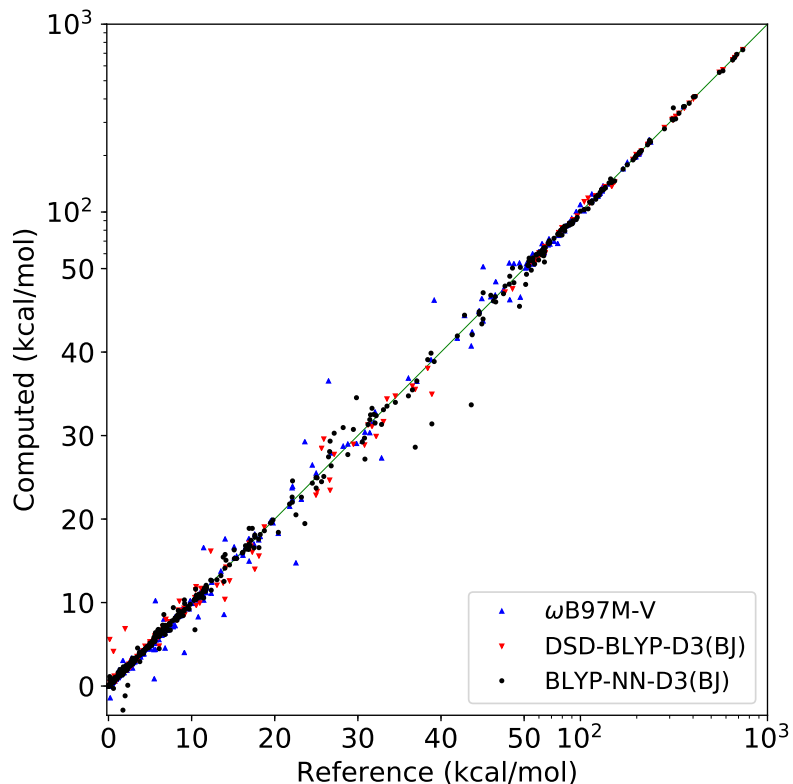


Figure 3.8: Error plot of the BLYP-based neural network functional, along with other methods.

are relatively stable. (The MAD becomes 2.37 kcal/mol, which is changed by 21%; the WTMAD becomes 2.72 kcal/mol, which is changed by only 6%.) The MAD is reliable when errors in similar reactions are averaged. It, however, can still ignore reactions with small reaction energies when different data sets are mixed. Therefore, we advocate using the WTMAD to properly report the quality of a method when it is tested on a mixture of various different data sets. As seen from the WTMAD values, neural network functionals can at least provide comparable accuracies to the best-performing hybrid and double-hybrid functionals, and in many cases actually outperform them. In Fig. 3.8 and Fig. 3.9, we provide error plots of neural network functionals, along with (double-)hybrid functionals.

While comparing neural network functionals that use different base functionals,

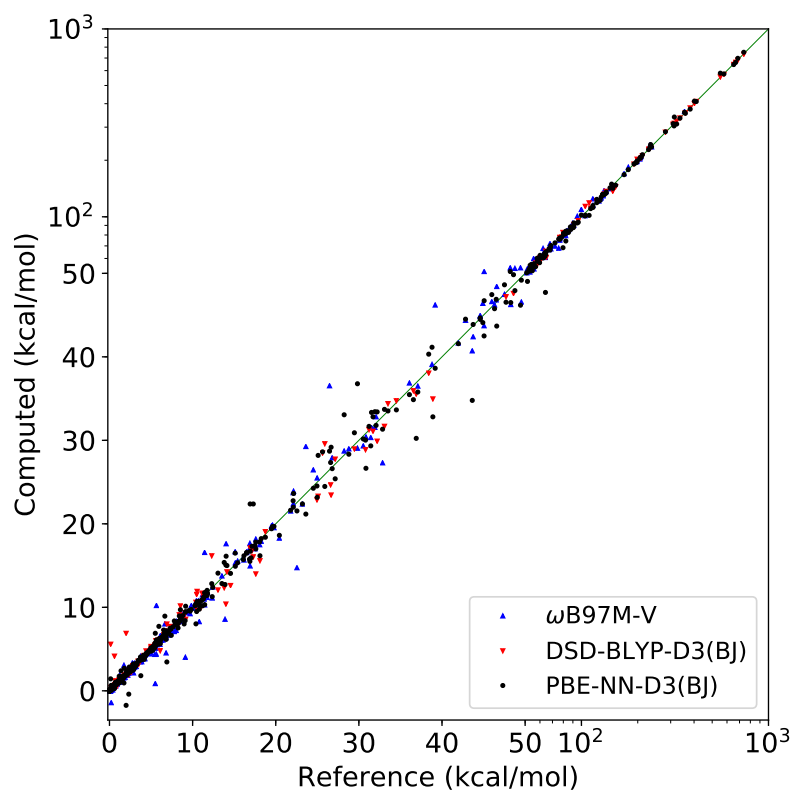


Figure 3.9: Error plot of the PBE-based neural network functional, along with other methods.

one can see that using BLYP as the base functional is generally better than PBE. The dependence on the base functional is expected, since we are effectively learning the error that the base functional is making, and the magnitude of the error can influence the quality of the atom-center expansion approximation.

Though this neural network functional is intended to be universal, we should be aware that its performance on a particular reaction is highly dependent on the training data fed to the model. We observed that the neural network might give some reaction energies that are worse than those provided by the base method, if the reactions contain elements that do not occur often in the training data. Even for elements that have high occurrence rates in the training data, a few large errors could also be made, such as the large error in the IP of the oxygen atom, as pointed out before. The reason for the bad performance is because, for electron densities in those systems, they are in regions that are far from where the neural network is trained for. Even if the neural network does not take the atomic number as an input directly, the local density is mostly determined by the element type. Therefore, before applying this method to systems containing elements that are currently not well explored by the training set used, it is necessary to first extend the training database with more data for those elements.

For the oxygen atom IP issue, one possible explanation is that, molecules containing oxygen included in the training set are mostly bonded molecules. The local density of an ionized O atom differs from what the neural network learned significantly, thus leading to a large error when extrapolating.

Next, we study how this functional behaves when faced with elements that are completely untrained. There are reactions in the GMTKN55 database that involve elements heavier than Kr. In addition, the MOR41 database can be used to benchmark the performance on transition metals. The details about the data sets can be

summarized as:

1. 30 non-covalent interaction energies from the HAL59 set. The untrained element is I.
2. The HEAVY28 set, which includes 28 non-covalent interaction energies that involve Sb, Te, I, Pb and Bi.
3. 4 reaction energies from the HEAVYSB11 set. The elements of interest are Sn, Sb, Te and Pb.
4. The MOR41 set, which includes 41 closed-shell organometallic reactions, and the metals in this set are all transition metals (from Ti to Pt).

It is interesting to know how the neural network functional performs on those reactions, and the impacts on the machine-learning model when part of these data are included in the training. Table 3.5 lists MADs on HAL59, HEAVY28 and HEAVYSB11 for base functionals, and neural networks that are: the model that is trained on light elements only; including 50% and 75% of above data in the training. It can be found that, without including any entries in those data sets in the training, BLYP-NN-D3(BJ) already outperforms BLYP-D3(BJ) in HAL59 and HEAVY28. This is because reaction energies in these two sets are all non-covalent interaction energies. The errors on untrained elements are mostly cancelled. Furthermore, including some of them in the training does improve the accuracy. On the contrary, it seems that PBE-NN-D3(BJ) does not benefit from the error cancellation. This is also reasonable because the set of DFT-D3 parameters may not be optimized. For molecules that contain heavy elements, they usually have stronger non-covalent interactions, the DFT-D3 parameters trained on light elements may give larger errors when applied on heavy elements. Fortunately, this problem can be solved by extending the

training database. The HEAVYSB11 set involves chemical reactions, thus the neural network functionals can no longer rely on error cancellation. This is also the case for the MOR41 set, which is presented in Table 3.6. Though including part of the data in the training still cannot make neural network functionals outperform their base methods, we can clearly observe a trend that, the quality of the neural network on transition metals are improved by extending the training data set.

In addition to the effect of the training database, the form of the loss function might also be worth studying. The WTMAD2 works best when similar reactions are grouped together. However, existing databases may have entries with quite different magnitudes within a subset. In a subset with a large mean reaction energy ($\overline{|\Delta E|}$), those reactions with small reaction energies will become insignificant when training. A particular example is HAT707MR_411 (reference value -4.69 kcal/mol, BLYP gives -4.59 kcal/mol, while the neural network gives -8.76 kcal/mol). In this case, other loss function forms such as the WTMAD1, might be more appropriate.

Finally, when it comes to study the fundamental difficulties that traditional DFAs have, which are the fractional charge error and the fractional spin error, because the neural network functional remains an explicit functional in the density, in theory it is still incapable of dealing with these difficulties. In Fig. 3.10 we show the fractional charge error for the carbon atom. It is interesting to note that the neural network functional behaves similarly as compared to regular semi-local functionals. It is possible that the local orbital scaling correction (LOSC) is able to further correct this neural network functional [97, 98].

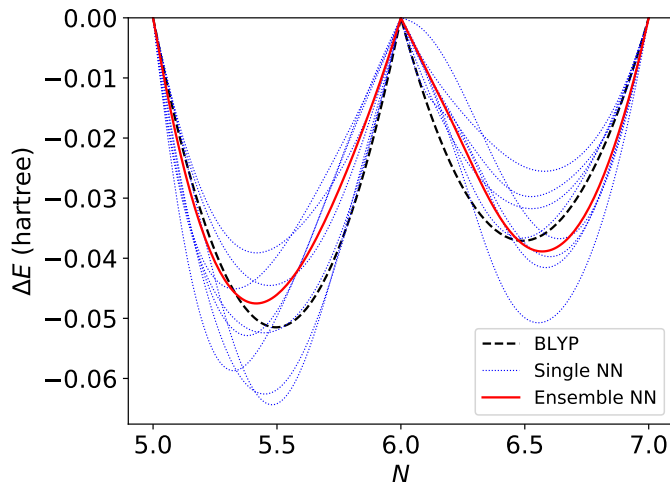


Figure 3.10: Fractional charge error of the carbon atom. It can be seen that the ensemble neural network is necessary here, as it smooths out the errors produced by single neural networks, making the final error have a shape close to what semi-local functionals usually have.

Table 3.4: Comparing neural network functionals with DSD-BLYP-D3(BJ) for tests belong to GMTKN55. Notations are the same as in Table 3.1. The unit is kcal/mol.

		MAD	RMSD	WTMAD
TC	BLYP-NN-D3(BJ)	2.99	6.49	2.89
	PBE-NN-D3(BJ)	3.57	6.29	3.76
	DSD-BLYP-D3(BJ)	1.74	2.66	3.28
BH	BLYP-NN-D3(BJ)	1.37	1.75	5.20
	PBE-NN-D3(BJ)	1.60	2.84	6.11
	DSD-BLYP-D3(BJ)	1.81	2.73	7.13
IE	BLYP-NN-D3(BJ)	1.48	2.94	4.79
	PBE-NN-D3(BJ)	1.41	2.35	5.86
	DSD-BLYP-D3(BJ)	1.06	1.87	4.44
NC	BLYP-NN-D3(BJ)	0.22	0.35	2.95
	PBE-NN-D3(BJ)	0.33	0.63	3.65
	DSD-BLYP-D3(BJ)	0.37	0.78	4.56

Table 3.5: Errors of 30 non-covalent interaction energies in the HAL59 set, and 28 non-covalent interaction energies in the HEAVY28 set. The unit is kcal/mol.

	HAL59			HEAVY28			HEAVYSB11
	MAD (100%)	MAD (50%)	MAD (25%)	MAD (100%)	MAD (50%)	MAD (25%)	MAD (100%)
BLYP-D3(BJ)	0.81	0.80	0.68	0.35	0.37	0.41	4.26
BLYP-NN-D3(BJ) (0% training)	0.30	0.35	0.26	0.19	0.14	0.17	5.20
BLYP-NN-D3(BJ) (50% training)	—	0.26	0.15	—	0.21	0.20	—
BLYP-NN-D3(BJ) (75% training)	—	—	0.15	—	—	0.28	—
PBE-D3(BJ)	1.09	1.10	0.96	0.38	0.35	0.40	3.60
PBE-NN-D3(BJ) (0% training)	1.90	1.88	1.65	0.67	0.58	0.50	6.27
PBE-NN-D3(BJ) (50% training)	—	0.33	0.19	—	0.76	0.26	—
PBE-NN-D3(BJ) (75% training)	—	—	0.15	—	—	0.19	—

Table 3.6: Errors in the MOR41 set. The unit is kcal/mol.

	MAD (100%)	MAD (50%)	MAD (25%)
BLYP-D3(BJ)	5.24	4.47	2.84
BLYP-NN-D3(BJ) (0% training)	6.58	6.80	5.56
BLYP-NN-D3(BJ) (50% training)	—	5.48	3.68
BLYP-NN-D3(BJ) (75% training)	—	—	3.38
PBE-D3(BJ)	3.84	3.52	3.17
PBE-NN-D3(BJ) (0% training)	8.46	9.55	8.33
PBE-NN-D3(BJ) (50% training)	—	5.75	3.82
PBE-NN-D3(BJ) (75% training)	—	—	4.63

Chapter 4

Improved Self-consistency for a Functional Tackles Delocalization Errors

This chapter is mainly adapted from the journal article:

Yuncai Mei, Zehua Chen, and Weitao Yang. “Self-Consistent Calculation of the Localized Orbital Scaling Correction for Correct Electron Densities and Energy-Level Alignments in Density Functional Theory.” *The Journal of Physical Chemistry Letters*, **2020**, *11*(23), 10269-10277.

4.1 Introduction

Density functional theory (DFT) [99, 100, 101] has been widely used to calculate and predict the electronic structure of molecular systems in practice. The performance of DFT depends on the quality of an applied density functional approximation (DFA) to the exchange-correlation energy E_{xc} . Although conventional DFAs, such as the local density approximation (LDA), [102, 103] general gradient approximations (GGAs), [104, 105, 106] or hybrid GGAs, [107, 108, 109] are commonly used in practice, they all have the delocalization error [4, 5, 6, 7] and fail to describe some critical problems.

The delocalization error in conventional DFAs [4, 110] exhibits a size-dependent manner as follows. (1) For systems with small numbers of atoms and a small physical extent, commonly used DFAs have good accuracy for describing the total energies of systems with integer numbers of electrons, but the delocalization error is exhibited as the convex deviation from the Perdew-Parr-Levy-Baldurzi (PPLB) linearity condition

for systems with fractional numbers of electrons [111, 112, 113]. (2) For systems with large numbers of atoms, commonly used DFAs have small errors for systems with fractional numbers of electrons and obey the fractional charge linearity condition at the bulk limit, but the delocalization error is manifest as large errors in total energies for systems with an integer number of electrons. (3) For systems with a finite number of atoms but with a large physical extent (as near a dissociation limit), the delocalization error leads to errors in the total energies of systems with integer electron numbers and produces the convex deviation from the PPLB condition for systems with fractional numbers of electrons. The cases of ionization energies for noniterating helium clusters clearly demonstrate these three scenarios. [110]

Such a delocalization error in conventional DFAs, shown as either the convex deviation to the PPLB condition (aforementioned point 1 and 3) or large errors in the total energies of systems with an integer number of electrons (aforementioned point 2), leads to a major error in the energy derivatives with respect to the electron number (i.e., the chemical potentials). [5, 114] Thus, this error leads the conventional DFAs to underestimate of the exact ionization potentials (IPs) from the HOMO (highest occupied molecular orbital) energies and the overestimation of electron affinities (EAs) from the LUMO (lowest unoccupied molecule orbital) energies. In addition to the chemical potential, the delocalization error would produce electron densities that are too delocalized, as the error falsely lowers the total energy of the system. [4, 6, 7] Examples reflecting this issue are the wrong dissociation limits of molecules [115, 116, 117, 118, 119] and the underestimation of reaction barriers [7] from conventional DFAs. For charge transfer complexes, the delocalization error leads to an overestimation of charge transfer and binding energies [6]. In terms of material interfaces and defects, the delocalization error can lead to incorrect charge transfer across the interfaces and significant errors in energy-level alignments [120, 121, 122, 123].

Interfaces and energy-level alignments play important roles in many technological applications: they strongly influence the charge extraction and transport in solar cell devices [124] and catalysis in semiconductors [125]. Thus, it remains an important challenge to describe the correct energy-level alignment for interfaces with DFAs.

To reduce the delocalization error, there has been much effort devoted to improving DFAs. These include the development of long-range corrected functionals [126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137] and double hybrid functionals [138, 139, 140, 141, 142]. It has been shown that these methods can reduce the delocalization error, but there remains a challenge for a systematic correction across system types, sizes, and scales.

Recently, our group developed the localized orbital scaling correction (LOSC) method, which imposes the PPLB condition by utilizing orbitalets (a set of molecular orbitals localized in both physical and energy spaces) to associated parent DFAs to reduce the delocalization error. [110, 143] Benefiting from the novel features of orbitalets that dynamically switch between the canonical orbitals (COs) and localized orbitals (LOs), correction from LOSC can be flexibly and automatically applied to the global or local region of the system. Therefore, the LOSC has the potential to reduce the delocalization error in a systematic and size-consistent way. It has been shown that LOSC greatly improves the prediction of quasiparticle energies – ionization energies and electron affinities from the eigenvalues of the generalized Kohn-Sham Hamiltonian. [110, 143, 144] Good performance from the LOSC has also been observed in ionized excited-state energies for atoms, small molecules, and very large systems. [145, 145] Furthermore, the LOSC can even outperform the many-electron Green’s function method of GW in some cases studied. [144]

In the original LOSC paper [110], two approaches for applying the LOSC to the parent DFAs have been discussed. One way is in the post-self-consistent field ap-

proach (post-LOSC) in which the converged electron density from the parent DFA (ρ_s^{DFA}) is directly used to evaluate the energy correction ΔE^{LOSC} from the LOSC. The other way is the self-consistent field approach (SCF-LOSC) in which the LOSC effective Hamiltonian $\Delta h^{\text{LOSC}} = \frac{\delta \Delta E^{\text{LOSC}}}{\delta \rho_s}$ is introduced into the DFA Hamiltonian h_s^{DFA} . After solving the KS-equations with the updated Hamiltonian $h_s = h_s^{\text{DFA}} + \Delta h^{\text{LOSC}}$ self-consistently, the converged electron density from LOSC-DFA ($\rho_s^{\text{LOSC-DFA}}$) is obtained and the correction ΔE^{LOSC} is then evaluated based on $\rho_s^{\text{LOSC-DFA}}$. In practice, the originally proposed SCF approach [110] for the SCF-LOSC calculation tends to encounter convergence problems easily, especially for large molecules, because only an approximate form of the LOSC effective Hamiltonian was used. Therefore, only the performance of post-LOSC has been well investigated. Although the post-LOSC has been demonstrated to show much improvement in the description of band gaps, total energies and photoemission spectra, [110, 143, 144] the development of a reliable SCF-LOSC approach is still necessary. The reasons are that the electron density of a molecular system is as important as the energy properties (total energies and orbital energies) since it is closely related to the molecule’s geometry, chemical bonding, and reaction reactivities. The conventional DFAs, suffering from the delocalization error, produce delocalized electron densities and underestimate the total energies in many cases. [115, 116, 117, 118, 119, 146] Applying post-LOSC in these cases is not sufficient because it improves only the energies for the parent DFA but leaves the significant error in electron density unchanged. Second, the post-LOSC, being a reasonable approximation to the SCF-LOSC, is valid only under the condition that ρ_s^{DFA} is close to $\rho_s^{\text{LOSC-DFA}}$. When ρ_s^{DFA} differs greatly from $\rho_s^{\text{LOSC-DFA}}$ as the delocalization error from the parent DFA becomes significant, the post-LOSC may not provide reliable results due to the lack of self-consistency, and a great difference between post-LOSC and SCF-LOSC can be expected.

4.2 Methods

In this work, we present a new and robust SCF approach to achieving reliable SCF-LOSC calculations. We start with reviewing the methodology of LOSC. LOSC is designed to impose the PPLB linearity condition locally on conventional DFAs by using orbitalets (LOs, $\{\phi_i\}$). The orbitalets are molecular orbitals localized in both physical and energy spaces and can be obtained through the LOSC localization procedure. Taking the localization in the developed version of LOSC (LOSC2) [143] as an example, the LOSC localization cost functional F takes the following form,

$$F(\rho_s, \{\psi_i\}, \mathbf{U}) = (1 - \gamma) \sum_p \left(\langle \mathbf{r}^2 \rangle_p - \langle \mathbf{r} \rangle_p^2 \right) + \gamma C \sum_p \left(\langle h[\rho_s]^2 \rangle_p - \langle h[\rho_s] \rangle_p^2 \right), \quad (4.1)$$

where

$$\langle X \rangle_p = \langle \phi_p | X | \phi_p \rangle, \quad X = \mathbf{r}, \mathbf{r}^2, h, h^2, \quad (4.2)$$

$$\text{and } \phi_p = \sum_i U_{pi} \psi_i, \quad (4.3)$$

\mathbf{U} is the unitary transformation matrix, h is the one-electron Hamiltonian of the associated DFA evaluated at the electron density ρ_s in each SCF cycle, and $\{\psi_i\}$ represents the corresponding COs. Through the defined localization, the obtained LOs are unitarily transformed from both occupied and virtual COs. On the basis of these LOs, the energy correction from the LOSC (ΔE^{LOSC}) is constructed to restore the linearity condition. ΔE^{LOSC} is expressed as

$$\Delta E^{\text{LOSC}} = \sum_{ij} \frac{1}{2} \kappa_{ij} \lambda_{ij} (\delta_{ij} - \lambda_{ij}), \quad (4.4)$$

in which $\boldsymbol{\kappa}$ is the LOSC curvature matrix and $\boldsymbol{\lambda}$ is the local occupation matrix. The LOSC curvature and local occupation matrices are directly generalized from the global scaling correction (GSC) method. [147] In contrast to GSC that constructs the curvature from COs and uses canonical occupation numbers to provide corrections globally, the LOSC evaluates the curvature based on LOs and uses local occupation numbers $\{\lambda_{ij}\}$ to impose the PPLB linearity condition locally. The LOSC curvature matrix has different expressions in different versions of LOSC. We call the original version of the LOSC LOSC1 [110] and the later version LOSC2 [143] in the following text. No matter which version of the LOSC is used, the LOSC curvature matrix is completely and explicitly determined by LOs. For example, the curvature matrix from LOSC1 is defined as

$$\kappa_{ij} = \int \frac{\rho_i(\mathbf{r})\rho_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - \frac{2\tau C_x}{3} \int [\rho_i(\mathbf{r})]^{\frac{2}{3}} [\rho_j(\mathbf{r})]^{\frac{2}{3}} d\mathbf{r}, \quad (4.5)$$

in which $\rho_i(\mathbf{r})$ is the local orbital's density and is defined as $\rho_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2$. The local occupation matrix λ is determined by LOs and the KS density operator ρ_s , and it is expressed as

$$\lambda_{ij} = \langle \phi_i | \rho_s | \phi_j \rangle. \quad (4.6)$$

With the defined LOs, LOSC curvatures and local occupation numbers, the original SCF procedure for SCF-LOSC [110] was proposed and the working flow is shown in Figure 4.1. As shown in Figure 4.1, one should notice that both LOs and the electron density are updated in every SCF cycle. The dependence of LOs on the electron density (as defined in Eq. 4.1) in each SCF cycle introduces complexities for deriving the exact LOSC effective Hamiltonian. Specifically, according to the original SCF procedure and Eqs. 4.4 - 4.6, the exact LOSC effective Hamiltonian Δh^{LOSC} for

the original SCF approach is given by

$$\Delta h^{\text{LOSC}} = \frac{\delta \Delta E[\{\phi_i\}, \rho_s]}{\delta \rho_s} = \frac{\delta \Delta E}{\delta \rho_s} \Big|_{\{\phi_i\}} + \sum_i \frac{\delta \Delta E}{\delta \phi_i} \Big|_{\rho_s} \frac{\delta \phi_i}{\delta \rho_s} + \sum_i \frac{\delta \Delta E}{\delta \phi_i^*} \Big|_{\rho_s} \frac{\delta \phi_i^*}{\delta \rho_s}. \quad (4.7)$$

The first term (denoted as Δh_1) on the r.h.s. of Eq. 4.7 is the explicit contribution with the LOs fixed, and the last two terms (denoted as Δh_2) are the implicit contribution from the relaxation of LOs because of the dependence on the electron density. Due to the difficulty of evaluating Δh_2 , this term was ignored in practical calculations. As shown in Figure 4.1, the LOSC effective Hamiltonian is approximated with only Δh_1 ,

$$\Delta h^{\text{LOSC}} \approx \Delta h_1 = \sum_i \kappa_{ii} \left(\frac{1}{2} - \lambda_{ii} \right) |\phi_i\rangle \langle \phi_i| - \sum_{i \neq j} \kappa_{ij} \lambda_{ij} |\phi_i\rangle \langle \phi_j|. \quad (4.8)$$

However, such an approximate LOSC effective Hamiltonian is not robust in practice, as we find it easily leads to nonconvergence. To solve the convergence problem, one straightforward solution is to derive the Δh_2 term and use the exact LOSC effective Hamiltonian. However, this would be complicated and difficult to achieve.

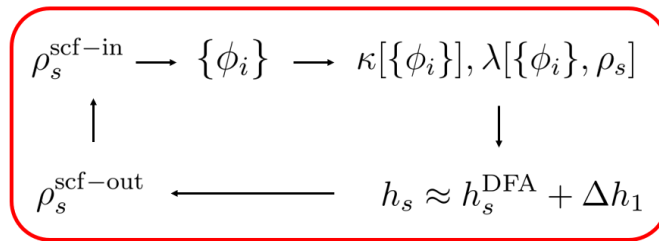


Figure 4.1: Original SCF procedure with the approximate LOSC effective Hamiltonian.

In this chapter, we develop an alternative solution to the problem. The idea is to define a new SCF procedure with the removal of LOs' dependence on the electron density in each SCF cycle. The key step is to redefine the LOSC localization proce-

dure, in other words, the localization cost functional. Instead of defining the LOSC localization cost functional that is dependent on the electron density and COs in each SCF cycle as shown in Eqs. 4.1 - 4.3, we can just use the electron density ρ_s^{DFA} and the corresponding COs $\{\psi_i^{\text{DFA}}\}$ from a converged DFA calculation in the cost functional. Therefore, we obtain a set of predetermined LOs $\{\phi_i^0\}$ in advance of the SCF-LOSC calculation and keep the same set of LOs $\{\phi_i^0\}$ during the SCF procedure. By applying this strategy to LOSC2 as an example, the modified localization cost functional is expressed as

$$F(\rho_s^{\text{DFA}}, \{\psi_i^{\text{DFA}}\}, \mathbf{U}) = (1 - \gamma) \sum_p \left(\langle \mathbf{r}^2 \rangle_p - \langle \mathbf{r} \rangle_p^2 \right) + \gamma C \sum_p \left(\langle h[\rho_s^{\text{DFA}}]^2 \rangle_p - \langle h[\rho_s^{\text{DFA}}] \rangle_p^2 \right), \quad (4.9)$$

where

$$\langle X \rangle_p = \langle \phi_p^0 | X | \phi_p^0 \rangle, \quad X = \mathbf{r}, \mathbf{r}^2, h, h^2, \quad (4.10)$$

$$\text{and } \phi_p^0 = \sum_i U_{pi} \psi_i^{\text{DFA}}, \quad (4.11)$$

h is now evaluated at ρ_s^{DFA} and LOs are unitarily transformed from the converged $\{\psi_i^{\text{DFA}}\}$. Clearly, the LOs in this new SCF approach do not depend on the electron density of each SCF cycle. This treatment of LOs in the new approach makes the Δh_2 term in Eq. 4.7 vanish and gives the exact LOSC Hamiltonian with the Δh_1 term only:

$$\Delta h^{\text{LOSC}} = \frac{\delta \Delta E[\{\phi_i^0\}, \rho_s]}{\delta \rho_s} = \sum_i \kappa_{ii} \left(\frac{1}{2} - \lambda_{ii} \right) |\phi_i^0\rangle \langle \phi_i^0| - \sum_{i \neq j} \kappa_{ij} \lambda_{ij} |\phi_i^0\rangle \langle \phi_j^0|. \quad (4.12)$$

The LOSC curvature is determined by the set of $\{\phi_i^0\}$, and it needs to be evaluated

only once. The local occupation matrix is evaluated based on $\{\phi_i^0\}$ and electron density during the SCF procedure: $\lambda_{ij} = \langle \phi_i^0 | \rho_s | \phi_j^0 \rangle$.

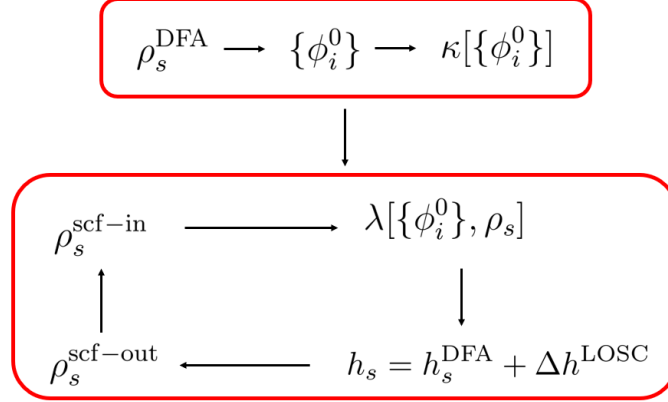


Figure 4.2: New SCF procedure with the exact LOSC Hamiltonian.

The workflow of the new SCF procedure for the SCF-LOSC calculation is demonstrated in Figure 4.2. It involves the following steps: (1) carry out the SCF convergence from the parent DFA to get the converged electron density ρ_s^{DFA} and COs $\{\psi_i^{\text{DFA}}\}$; (2) apply the LOSC localization procedure to generate $\{\phi_i^0\}$, based on ρ_s^{DFA} and $\{\psi_i^{\text{DFA}}\}$; (3) construct and store the curvature matrix evaluated from $\{\phi_i^0\}$; (4) use ρ_s^{DFA} as the initial guess to start the SCF-LOSC calculation associated with parent DFA; (5) use the density ρ_s from current SCF cycle to construct the DFA Hamiltonian h_s^{DFA} ; (6) evaluate the local occupation matrix based on ρ_s and $\{\phi_i^0\}$ and construct the exact LOSC effective Hamiltonian Δh^{LOSC} via Eq. 4.12; (7) apply Δh^{LOSC} to h_s^{DFA} and update COs and ρ_s ; (8) check the convergence and go back to step (5) if it is not converged.

In the new SCF-LOSC approach, since the LOSC curvature matrix needs to be evaluated only once and updating the local occupation, $\lambda_{ij} = \langle \phi_i^0 | \rho_s | \phi_j^0 \rangle$, is simple, the computational cost for the new SCF-LOSC approach is only about 2 times that of the conventional DFA-SCF calculation. Specifically, one is the generation of LOs

from a one-time conventional DFA-SCF calculation, and the other is the SCF-LOSC calculation with the corrected KS Hamiltonian, with fixed LOs and an LOSC curvature matrix.

Comparing the new SCF approach with the original one, the solution from the new SCF approach may be different from the original SCF-LOSC solution. This is because, at the SCF solution point, the LOs used to evaluate the total energy in the new SCF-LOSC approach are always obtained from ρ_s^{DFA} rather than $\rho_s^{\text{LOSC-DFA}}$. The significance of this difference needs to be verified with numerical results. If the relaxation of LOs, like in the original SCF-LOSC, turns out to be necessary, then we can apply an additional layer of the SCF cycle on top of the new SCF-LOSC procedure in order to update the LOs. This two-layer SCF method is noted as macro-SCF-LOSC.

4.2.1 Macro-SCF-LOSC

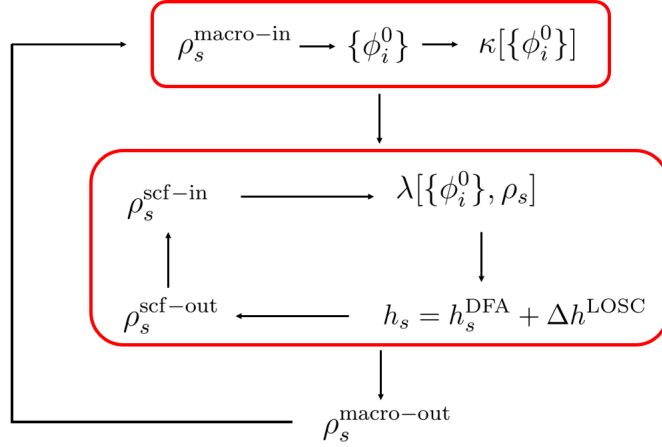


Figure 4.3: The macro-SCF procedure for LOSC.

The workflow of the macro-SCF-LOSC is shown in Figure 4.3. The macro-SCF-LOSC includes two layers of SCF cycles. The outer cycle is to update the LOs. The

inner cycle is to update the electron density with the given LOs. The inner cycle shares the same spirit of the new SCF-LOSC approach. In details, the macro-SCF-LOSC involves the following steps:

1. start the macro-SCF-LOSC with an initial density guess $\rho_s^{\text{macro-in}}$;
2. apply LOSC localization procedure to produce LOs $\{\phi_i^0\}$ based on $\rho_s^{\text{macro-in}}$;
3. construct LOSC curvature matrix from $\{\phi_i^0\}$;
4. achieve the convergence of the inner SCF-LOSC calculation with the given $\{\phi_i^0\}$ and $\kappa[\{\phi_i^0\}]$ and obtain the converged density as $\rho_s^{\text{macro-out}}$;
5. check the convergence of the macro-SCF-LOSC; go back to step 2 if it is not converged.

Ideally, because the macro-SCF-LOSC optimizes both electron density and LOs, the macro-SCF-LOSC should yield the same results as the original SCF-LOSC approach. From the numerical results shown in the following text, we find that the macro-SCF-LOSC is not necessary in practice because the new SCF approach without macro iterations is already able to provide excellent reliable results.

4.3 Computational Details

All the DFT calculation are performed from using an in-house developed QM⁴D package [38], if not specified. The MRCI+Q calculation is performed from using GAMESS package. [148, 149] The GW calculation is performed from using FHI-aims package. [150, 151] The MP2 and the D3 version of Grimme’s dispersion with Becke-Johnson damping (D3BJ) are performed from using Gaussian 16 package [95].

cc-pVTZ is used as the basis set for all the calculations. For LOSC calculations, aug-cc-pVTZ-RIFIT is used as the fitting basis set in the construction of LOSC curvature matrices. The molecular geometries for 1,4-benzenediamine and tetracyanoethylene (TCNE) are optimized from using B3LYP/6-31g* with Gaussian 16 package. [95] In the following text, the D-A complex refers to the 1,4-benzenediamine/TCNE complex system. The donor molecule refers to 1,4-benzenediamine and the acceptor refers to TCNE.

We use the same test sets to investigate the performance of the new SCF-LOSC approach for the atomization energies, reaction barriers, IPs and EAs. See Ref 143 for more information about molecular geometries and reference values for these test sets. All the DFT calculation are performed from using an in-house developed QM⁴D package. [38] 6-311++G(3df, 3pd) is used as the basis set for most of the results, while cc-pVDZ is used as the basis set for polyacetylene calculations. Aug-cc-pVTZ-RIFIT is the fitting basis used in the construction of curvature matrix in LOSC.

4.4 Results

Now, we check the performance of the new SCF approach for SCF-LOSC calculations. In the following text, the results from SCF-LOSC refer to the new SCF approach, if not specified otherwise. We first study the SCF convergence by testing a long organic molecule, the polyacetylene with 9 units of monomer ($[\text{C}_2\text{H}_2]_9$). As shown in Figure 4.4, the original SCF approach with the approximate LOSC effective Hamiltonian shows oscillations in the total energy along iterative steps, and it cannot reach convergence after about 20 steps. The new SCF procedure with the exact LOSC effective Hamiltonian converges smoothly and quickly.

Next, we study the effect of the initial guess in the new SCF approach. Note that

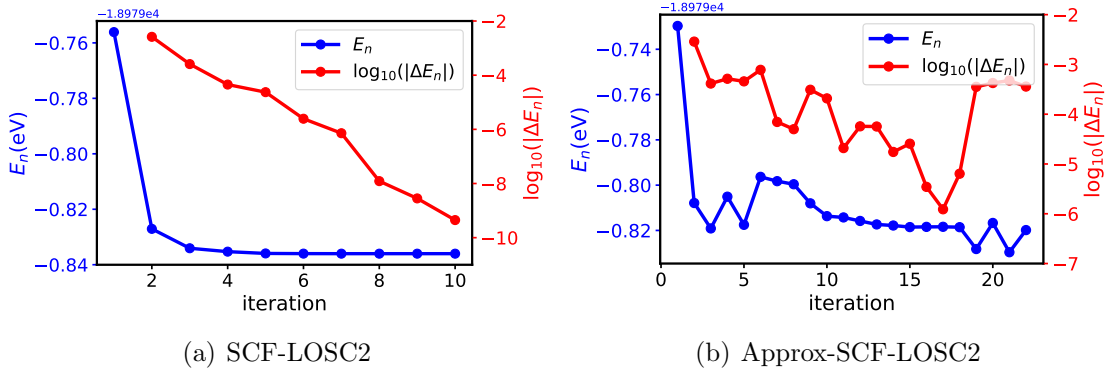


Figure 4.4: Comparison of SCF performance between (a) SCF-LOSC2 and (b) the original SCF-LOSC2 with an approximate Hamiltonian (approx-SCF-LOSC2) for polyacetylene ($n = 9$). E_n is the total energy at the n -th iteration step, and $\Delta E_n = E_n - E_{n-1}$. BLYP is used as the parent functional. 6-31g* is used as the basis set. aug-cc-pVTZ-RIFIT is used as the density fitting basis in LOSC2 curvature matrix construction.

in step (4) of the new SCF procedure we use ρ_s^{DFA} as the SCF initial guess. This is necessary because the LO geometric orientation is fixed in the SCF process. Using ρ_s^{DFA} as the initial guess ensures that the orientation of the LOs agrees with the initial density and orientation of COs at the beginning. If the used initial electron density has a different orientation from that of the LOs, then it will artificially produce fractional local occupation numbers in the initial step, which may cause the SCF calculation to converge to unphysical states with wrong energies. To support the discussion, Table 4.1 shows results for the test on the F atom, with a partially filled p shell. For such a small system, the LOs from LOSC localization are just the converged COs from the parent DFA. In other words, the localization is not operative and orbitals are just the COs for small systems. As shown in Table 4.1, if we use the nonrotated ρ_s^{DFA} as the initial guess, then the local occupation numbers will be exact integers (either 1 or 0), and LOSC (both LOSC1 and LOSC2) gives zero correction to the total energy of the parent DFA. However, if we use the rotated ρ_s^{DFA} as the initial guess, which has a set of rotated p orbitals, then the fractional local occupation

Table 4.1: Testing on the F atom for the effect of the orientation of initial guess to the new SCF-LOSC approach. BLYP is used as the parent functional. The converged densities from BLYP (rotated/non-rotated) are used as the initial guess. cc-pVTZ is used the basis set. aug-cc-pVTZ-RIFIT is used as the density fitting basis set in LOSC curvature matrix construction. Grid type is (99, 590).

Method	Total energy	energy difference	correction (initial) ³	correction (SCF) ⁴
BLYP	-99.7522856431			
SCF-LOSC1: non-rotated ¹	-99.7522856430	6.31×10^{-11}	1.10×10^{-14}	1.30×10^{-14}
SCF-LOSC1: rotated ²	-99.7794002845	-2.71×10^{-02}	-2.68×10^{-02}	-2.71×10^{-02}
SCF-LOSC2: non-rotated ¹	-99.7522856430	6.31×10^{-11}	1.00×10^{-14}	8.00×10^{-15}
SCF-LOSC2: rotated ²	-99.7522340746	5.16×10^{-05}	5.27×10^{-05}	5.05×10^{-05}

¹The orientation of the initial electron density matches that of LOs. ²The orientation of the initial electron density does not match that of LOs. ³LOSC energy correction for the first SCF cycle. ⁴LOSC energy correction for the last SCF cycle.

numbers will appear artificially and LOSC gives a nonzero correction to the total energy. In the case of LOSC1, we can see that such an artificial correction is even negative and leads to an unphysical state with an energy even lower than that of the DFA. In the case of LOSC2, the artificial energy correction is much smaller. This is because LOSC2 preserves the symmetry better than LOSC1 [143], causing these artificial energy corrections from each fractional local occupation number to almost cancel each other. As a whole, to avoid issues introduced by the orientation of the initial guess, we use ρ_s^{DFA} as the initial guess for the new SCF approach. Note that the choices of LO orientations will not be an issue at all, if the macro-SCF-LOSC is used because the LO will be generated from $\rho_s^{\text{LOSC-DFA}}$ in the macroiterations.

To compare the performance between new SCF-LOSC and the macro-SCF-LOSC, Table 4.2 and 4.3 show the total energies and Mulliken charges from these two methods for stretched HF, LiH and LiF molecules at large distance. In the cases of stretched HF and LiH, in which the parent functional B3LYP already describes the electron density well, the results from SCF-LOSC and macro-SCF-LOSC are almost the same. In the case of stretched LiF, in which the parent functional B3LYP yields

Table 4.2: Total energies (in a.u.) comparison between SCF-LOSC2 and macro-SCF-LOSC2. LOSC2 calculation is associated with B3LYP functional. LiH and LiF molecules are stretched at 14 bohr, and HF molecule is stretched at 16 bohr.

Molecule	SCF-LOSC2	macro-SCF-LOSC2	difference
HF	-100.2655352	-100.2655352	-1.01×10^{-12}
LiH	-7.994964751	-7.994964802	-5.06×10^{-08}
LiF	-107.2563249	-107.2565187	-1.94×10^{-04}

Table 4.3: Mulliken charge (in a.u.) comparison between SCF-LOSC2 and macro-SCF-LOSC2. LOSC2 calculation is associated with B3LYP functional. LiH and LiF molecules are stretched at 14 bohr, and HF molecule is stretched at 16 bohr. The Mulliken charge refers to F atom in HF, H atom in LiH and F atom in LiF.

Molecule	SCF-LOSC	macro-SCF-LOSC	difference
HF	0.000000	0.000000	0.00
LiH	0.000313	0.000311	-2.00×10^{-06}
LiF	0.000310	0.000240	-7.00×10^{-05}

delocalized electron density, SCF-LOSC still provide similar results to the ones from macro-SCF-LOSC with the difference of total energy only up to $1e-4$ a.u. (< 0.1 kcal/mol). Those results indicate the results from the new SCF-LOSC is reliable. Also it suggests that applying macro-SCF-LOSC would not be necessary in practice.

With the smooth convergence from the new SCF procedure, we test the performance of the new SCF-LOSC with the same test sets used in the development of LOSC, which include atomization energies, reaction barriers, first ionization potentials (IPs) and electron affinities (EAs). Table 4.4 shows the results from different methods for atomization energies, reaction barriers, IPs and EAs. The negative HOMO/LUMO energies from DFT calculation is used to evaluate the first IP/EA respectively. Table 4.5 shows the comparison of the negative HOMO energies of polyacetylene from different methods with the first IP.

Table 4.4: Mean absolute errors (MAEs) for atomization energies (AE), reaction barriers (RB), IPs and EAs from different methods. Results of AE and RB test sets is in kcal/mol. Results of IP and EA test sets is in eV.

Test set	DFA			post-LOSC2			SCF-LOSC2		
	LDA	BLYP	B3LYP	LDA	BLYP	B3LYP	LDA	BLYP	B3LYP
AE									
G2-1	42.20	4.94	2.45	42.19	4.93	2.45	42.19	4.93	2.45
NonHydrocarbon	97.74	14.73	7.64	97.74	14.73	7.65	97.74	14.73	7.65
Hydrocarbon	140.95	10.36	3.47	140.94	10.38	3.48	140.94	10.38	3.48
SubHydrocarbon	124.34	6.23	2.52	124.32	6.24	2.53	124.32	6.24	2.52
Radical	78.02	3.89	2.26	78.00	3.91	2.26	78.00	3.91	2.26
RB									
HTBH38	17.36	7.67	4.35	17.17	7.49	4.35	17.21	7.54	4.35
NHTBH38	14.22	10.11	6.38	14.24	10.14	6.38	14.23	10.12	6.38
IP & EA									
IP	4.29	4.50	3.19	0.49	0.62	0.35	0.50	0.62	0.35
EA	3.79	3.41	2.57	0.53	0.50	0.44	0.54	0.51	0.44

Table 4.5: Negative HOMO energies (in eV) of polyacetylene with different unit numbers calculated from different DFAs and compared with the first IP from RASPT2.

units	RASPT2	DFA			post-LOSC2			SCF-LOSC2		
		LDA	BLYP	B3LYP	LDA	BLYP	B3LYP	LDA	BLYP	B3LYP
1	10.48	6.71	6.34	7.46	10.46	10.11	10.46	10.46	10.11	10.46
2	9.18	5.84	5.46	6.43	9.49	9.12	8.35	9.49	9.12	8.35
3	8.18	5.40	5.01	5.89	8.84	8.47	8.03	8.84	8.47	8.03
4	7.69	5.14	4.74	5.56	8.01	7.66	7.77	8.01	7.66	7.77
5	7.33	4.97	4.56	5.34	7.78	7.41	7.59	7.80	7.42	7.60
6	7.04	4.84	4.43	5.17	7.56	7.19	7.41	7.59	7.21	7.42
7	6.85	4.75	4.33	5.05	7.48	7.10	7.19	7.53	7.15	7.20
8	6.66	4.68	4.25	4.95	7.45	6.95	7.15	7.53	7.03	7.18
9	6.56	4.62	4.19	4.87	7.23	6.85	6.93	7.32	6.94	6.97
10	6.41	4.57	4.14	4.80	7.08	6.61	6.84	7.19	6.72	6.89
MAE		2.49	2.89	2.09	0.50	0.20	0.33	0.54	0.24	0.35

In general, the new SCF-LOSC can be conducted easily in these test sets. For the test sets related to atomization energies and reaction barriers, both SCF-LOSC and post-LOSC currently preserve the performance as DFAs. For example, the mean absolute error (MAE) of G2-1 test set for atomization energies is 4.94 kcal/mol for BLYP, and 4.93 kcal/mol for both post-LOSC2-BLYP and SCF-LOSC2-BLYP. The MAE of HTBH38 test case for the reaction barriers is 7.67 kcal/mol for BLYP, 7.49 kcal/mol for post-LOSC2-BLYP, and 7.54 kcal/mol for SCF-LOSC2-BLYP. Such a performance is expected because in most of cases these tests are used with small molecular sizes and large orbital energy gaps between HOMO and LUMO, which makes orbitalets equal to the COs of the parent DFA and causes the local occupation matrix to be diagonal with integer numbers (1 for occupied space and 0 for virtual space). According to energy correction from LOSC shown in Eq. 4.4, these integer local occupation numbers give zero correction to the total energies. Retuning the parameters in the LOSC localization cost function to obtain more balanced localization between the physical space and energy space should provide a better performance. [143] However, this task is beyond the scope of the current work and will be studied in the future. For the test sets related to IPs and EAs, the new SCF-LOSC improves upon parent DFAs, and performs very similarly to the post-LOSC for these tested cases. For example, the MAE of IP test set is 4.50 eV for BLYP, and 0.62 eV for both post-LOSC2-BLYP and SCF-LOSC2-BLYP.

In the following text, we mainly focus on presenting the results that can be significantly different with and without self-consistency. They are all related to electron densities and energy levels (quasiparticle energies) associated with the molecular binding/dissociation processes. We first investigate the dissociation of three diatomic molecules (LiF, LiH and HF). B3LYP [104, 105, 152] is used as the parent functional because the SCF convergence from B3LYP can be easily reached at long bond dis-

tances for these molecules. Similar GGA calculations would show even larger delocalization error with charge density but was not obtained because of the failure in SCF for long bond lengths. Results from the multireference configuration interaction method with the Davidson correction (MRCI+Q) [153, 154, 155] are used as the reference. To study the description of the electron density, we look at Mulliken charges from Mulliken population analysis [156] and relative total energies of the molecules to their dissociation limits.

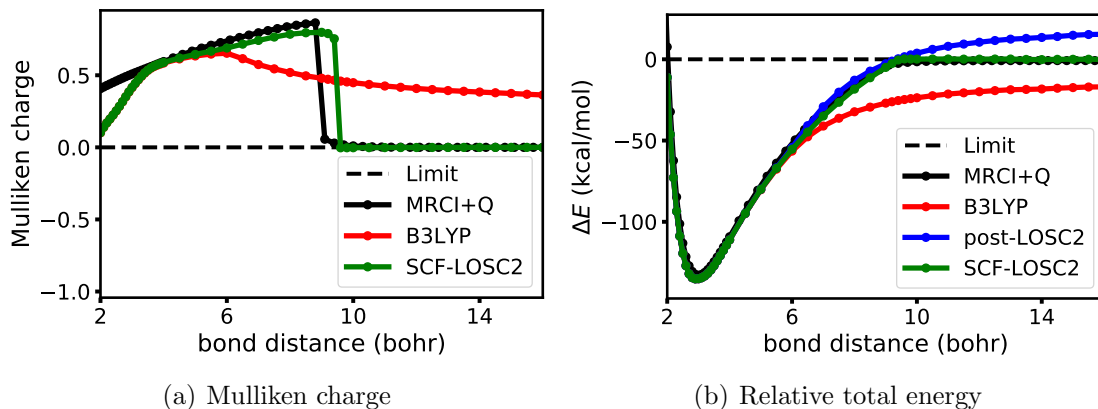


Figure 4.5: Dissociation of the LiF molecule: (a) Mulliken charge on the Li atom and (b) relative total energy of LiF with respect to Li and F atoms, $\Delta E = E_{\text{LiF}} - (E_{\text{Li}} + E_{\text{F}})$. LOSC2 calculations are associated with B3LYP.

Among the three diatomic molecules, B3LYP shows obvious delocalization errors for LiF and LiH, making them good cases for testing the performance of the new SCF-LOSC approach. The results for LiF are shown in Figure 4.5. Because $IP_{\text{Li}} > EA_{\text{F}}$, the LiF molecule must dissociate into neutral Li and F atoms (as with all neutral diatomic molecules). Clearly, according to Figure 4.5, we see B3LYP shows a significant delocalization error in electron density, which is reflected by the positive Mulliken charge of the Li atom and underestimated dissociation energy at the dissociation limit. On the basis of the delocalized electron density from B3LYP, post-LOSC-B3LYP corrects the total energy too much and yields a higher dissocia-

tion limit. In contrast to post-LOSC-B3LYP, SCF-LOSC-B3LYP corrects the electron density. As shown in Figure 4.5, the Mulliken charges from SCF-LOSC-B3LYP match well with the MRCI+Q reference. In addition, the relative total energies from SCF-LOSC-B3LYP almost overlap with the reference and shows the correct dissociation limit.

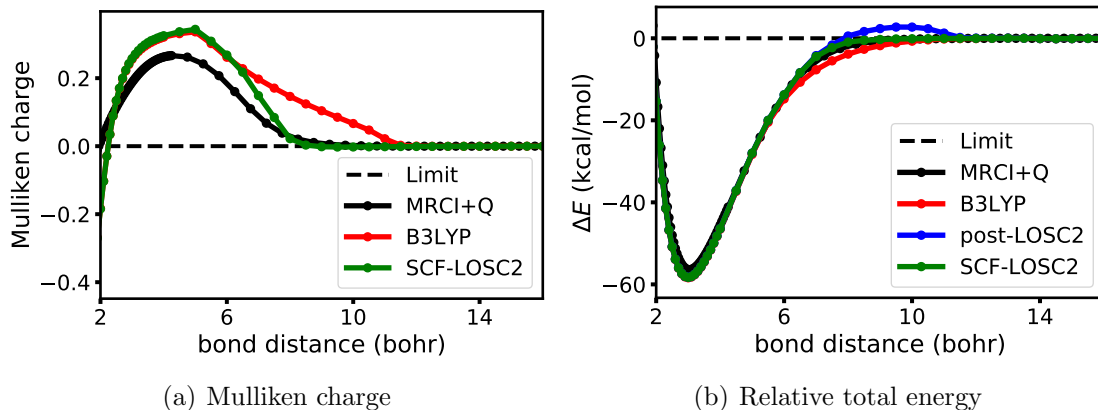


Figure 4.6: Dissociation of the LiH molecule: (a) Mulliken charge on the Li atom and (b) relative total energy of LiH with respect to Li and H atoms, $\Delta E = E_{\text{LiH}} - (E_{\text{Li}} + E_{\text{H}})$. LOSC2 calculations are associated with B3LYP.

Next, we look at LiH as shown in Figure 4.6. Because $\text{IP}_{\text{Li}} > \text{EA}_{\text{H}}$, the LiH molecule must dissociate into neutral Li and H atoms. According to Figure 4.6, we notice that, although B3LYP gives the correct Mulliken charge (zero charge) for the Li atom at the dissociation limit, it shows delocalization errors in the range of around 8-10 bohrs. MRCI+Q gives almost zero Mulliken charges in this range, while B3LYP gives positive charges. Such delocalized electron densities (around 8-10 bohrs) make the relative total energy from B3LYP lower than the MRCI+Q reference. In addition, errors in electron densities from B3LYP in this range lead the post-LOSC-B3LYP to yield relative total energies that are too high, shown as the small bump in the dissociation energy curve in Figure 4.6. From the results of SCF-LOSC-B3LYP, we observe that electron densities are corrected, shown as the Mulliken charges on the

Li atom in the range are corrected down to zero. With corrected electron densities from SCF-LOSC-B3LYP, the relative total energies from SCF-LOSC-B3LYP match much better with the MRCI+Q reference than B3LYP and post-LOSC-B3LYP.

Figure 4.7 shows the dissociation results for HF molecule. Since $IP_H > EA_F$, the HF molecule must dissociate into neutral H and F atom. According to Figure 4.7, we see that B3LYP already describes the electron density quite well for this system. In addition, we see SCF-LOSC-B3LYP preserves the good results for the Mulliken charge, and both SCF-LOSC-B3LYP and post-LOSC-B3LYP preserve the good results for the relative total energy.

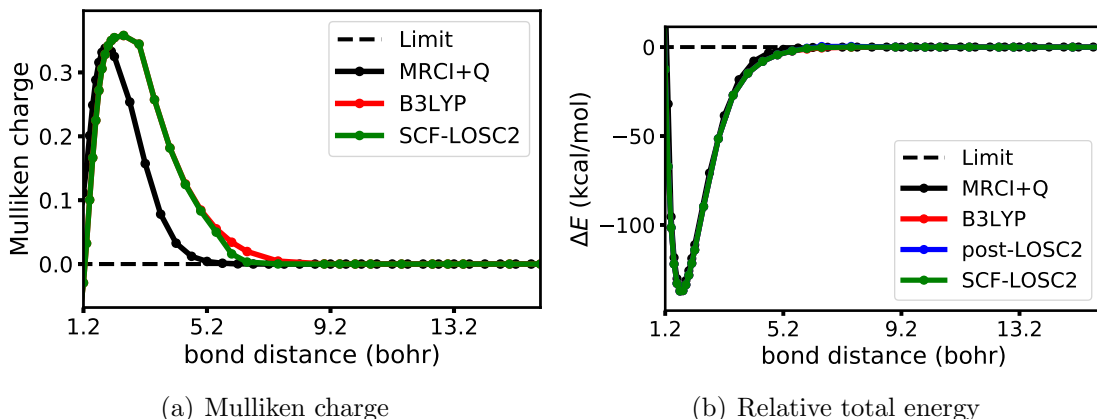


Figure 4.7: Dissociation of the HF molecule: (a) Mulliken charge on the H atom and (b) relative total energy of HF with respect to H and F atoms. $\Delta E = E_{HF} - (E_H + E_F)$. The LOSC2 calculation is associated with B3LYP.

We also study a donor-acceptor (D-A) organic complex system to demonstrate the performance of new SCF-LOSC for more complicated and larger systems. The donor molecule is 1,4-benzenediamine and the acceptor molecule is tetracyanoethylene (TCNE). Because $IP_D > EA_A$, the D-A complex will dissociate into two neutral subsystems. Figure 4.8 shows Mulliken charges and dissociation energies from DFT and MP2 [158] calculations. BLYP is used as the parent functional. It is clear that BLYP shows a great delocalization error at the dissociation limit. The donor molecule

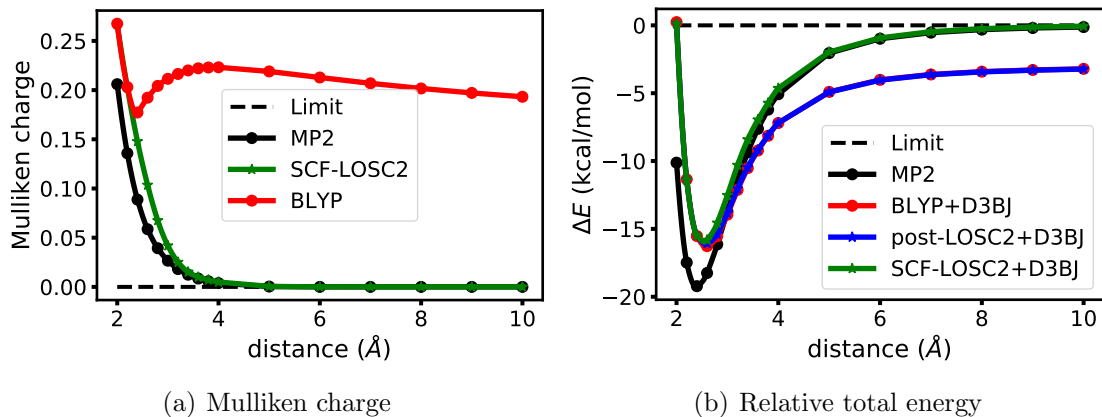


Figure 4.8: Dissociation of the donor-acceptor (D-A) complex (donor: 1,4-benzene-diamine, acceptor: tetracyanoethylene (TCNE)) from different methods: (a) Mulliken charge on the donor molecule and (b) relative total energy of the D-A complex with respect to neutral donor and acceptor molecules, $\Delta E = E_{DA} - (E_D + E_A)$. LOOSC2 calculations are associated with BLYP. The D3 version of Grimme’s dispersion with Becke-Johnson damping (D3BJ) [157] from the BLYP functional is added to all of the DFT energy results.

has a spurious positive charge, which means there is partial charge transfer from the donor to the acceptor molecule and the electron density is delocalized incorrectly. Due to the delocalized electron density from BLYP, the dissociation energies from BLYP and post-LOSC-BLYP show similar errors. With the correction to the electron density, we see that applying the new SCF-LOSC gives correct Mulliken charges and total energies along the dissociation coordinates, which include the dissociation limit.

In addition to the Mulliken charge analysis and dissociation energy of this D-A complex system, we examine its energy-level alignment (the first IP and EA) along the binding distance. Figure 4.9 shows the trend of first IP and EA with respect to the change of separation distance. The experimental IP of the donor and EA of the acceptor molecule are plotted as the reference values for the dissociation limit. Along the binding distance, the D-A complex system is calculated with DFT and GW methods (G_0W_0 and eigenvalue self-consistent GW (evGW)). For the DFT calculations, the negative orbital energy of HOMO and the negative orbital energy of LUMO are

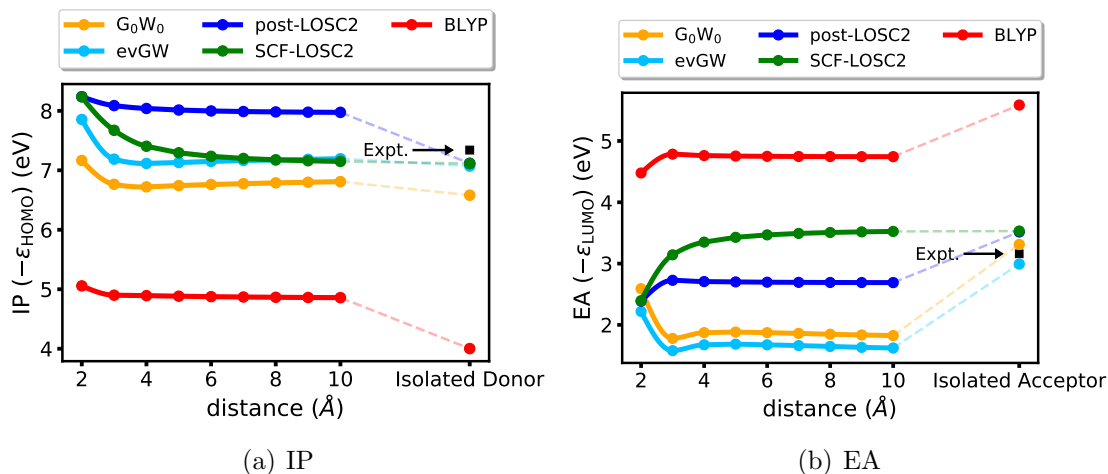


Figure 4.9: HOMO and LUMO energy-level alignment of the donor-acceptor (D-A) complex (donor: 1,4-benzenediamine, acceptor: tetracyanoethylene (TCNE)) from different methods: (a) First IP ($-\epsilon_{\text{HOMO}}$) of the D-A complex along the separation distance from 2 to 10 Å. At the right end of the figure, the first IP of the isolated donor molecule is plotted to be compared with the first IP from D-A complex calculations. The experimental IP of the isolated donor molecule is 7.34 eV [159] and marked with an arrow. (b) First EA ($-\epsilon_{\text{LUMO}}$) of the D-A complex along the separation distance from 2 to 10 Å. At the right end of the figure, the first EA of the isolated acceptor molecule is plotted to be compared with the first EA from D-A complex calculations. The experimental EA of the isolated acceptor molecule is 3.16 eV [160] and marked with an arrow. LOSC2 and GW calculations are based on BLYP.

used to evaluate the first IP and EA respectively [161]. For GW calculations, the obtained quasiparticle energies are used to evaluate the first IP and EA accordingly. According to Figure 4.9, we see BLYP shows significant underestimation of the IP and overestimation of the EA because of the delocalization error. The results from the GW method, which is based on BLYP, are also affected by the error in the electron density from BLYP. Especially in the case of the EA, the results from GW show an obvious underestimation. At large distances, even at the level of evGW, the error in the EA energy level is underestimated by about 1.5 eV, which is significant. Since SCF-LOSC-BLYP corrects the electron density, the results from SCF-LOSC-BLYP are close to the reference value at the dissociation limit, especially for the first IP. The results from post-LOSC-BLYP show much improvement as compared to BLYP but are not as good as those from SCF-LOSC-BLYP.

The description of the interface charge distribution and energy-level alignment in this charge transfer system demonstrates the major improvement from the SCF-LOSC and clearly highlights the importance of getting correct density distributions through self-consistent calculations for the correct energy-level alignment in DFT as well as in Green's function calculations.

Chapter 5

Conclusions

We present a new DFT approach in Chapter 2. It uses the (G)KS description for the auxiliary systems, but for the physical system, the energy functional is a major departure from the (G)KS energy functional form. The utilization of generalized auxiliary systems, combined with the linear response theory, leads to natural and rigorous multireference energy functionals. Currently the GOEP method is used to obtain self-consistency, and the pp-RPA and the SF-TDDFT are shown as examples. Numerical results show the capability of our method for describing systems with multireference character such as single bond breaking and double bond rotation. Significant improvement in the energy profile, including relative energies and equilibrium geometries, is observed when comparing this method with corresponding parent methods. Excitation energies are also evaluated. As a method that describes the ground state and excited states on the same footing, it delivers better or comparable accuracy for excitation energies compared to the non-self-consistent pp-RPA and the SF-TDDFT. In order to explore more possibilities on reference functionals, a new and better DFT functional, such as the LOSC[110, 143] may be applied. This work opens promising pathways to describe static/strong correlation within the density functional theory framework.

We have developed an energy correction to existing semi-local functionals in Chapter 3. This correction takes the electron density as the input, and is expressed as a summation of atomic contributions. The decomposed correction is then fitted by artificial neural networks. Tests have shown that, with GGAs (BLYP and PBE) as

the base method, this neural network functional can achieve a level of accuracy that is close to double-hybrid functionals, while keeps the method computationally affordable. Meanwhile, the tests on heavy elements and transition metals show that this method is possible to be further improved by extending the training data set with more accurate reference data. This opens up a new direction on the development of non-local functionals, and functionals of this type may help to generate more accurate data for large chemical systems that are important in the drug design or material design, where accurate wave-function methods or double-hybrid functionals can be too expensive to be applied.

The new SCF-LOSC approach presented in Chapter 4 overcomes the convergence issue and is very effective in practice. More importantly, we observe that the new SCF-LOSC approach is able to produce the correct electron densities, total energies, and energy-level alignments. The performance of SCF-LOSC is more reliable than that of post-LOSC, especially for the cases where converged densities from the parent DFA show significant delocalization errors. With the good performance and reliability, we believe that the new SCF-LOSC method will be promising for studying problems related to electron densities, and quasiparticle energy-level alignments in large molecules and interface systems.

Bibliography

- [1] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3B):B864, 1964.
- [2] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133, 1965.
- [3] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 76(12):6062–6065, 1979.
- [4] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Localization and Delocalization Errors in Density Functional Theory and Implications for Band-Gap Prediction. *Phys. Rev. Lett.*, 100(14):146401, April 2008. doi: 10.1103/PhysRevLett.100.146401.
- [5] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Fractional charge perspective on the band gap in density-functional theory. *Phys. Rev. B*, 77(11):115123, March 2008. doi: 10.1103/PhysRevB.77.115123.
- [6] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into Current Limitations of Density Functional Theory. *Science*, 321(5890):792–794, August 2008. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1158722.
- [7] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for Density Functional Theory. *Chem. Rev.*, 112(1):289–320, January 2012. ISSN 0009-2665. doi: 10.1021/cr200107z.
- [8] Weitao Yang and Qin Wu. Direct method for optimized effective potentials in density-functional theory. *Physical Review Letters*, 89(14):143002, 2002.
- [9] Qin Wu and Weitao Yang. A direct optimization method for calculating density functionals and exchange–correlation potentials from electron densities. *The Journal of Chemical Physics*, 118(6):2498–2509, 2003.
- [10] Eberhard KU Gross and Reiner M Dreizler. *Density Functional Theory*, volume 337. Springer Science & Business Media, 2013.
- [11] Ulf von Barth and Lars Hedin. A local exchange-correlation potential for the spin polarized case. i. *Journal of Physics C: Solid State Physics*, 5(13):1629, 1972.

- [12] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.
- [13] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098, 1988.
- [14] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the collesalvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785, 1988.
- [15] John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244, 1992.
- [16] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865, 1996.
- [17] Axel D Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.
- [18] Axel D Becke. A new mixing of hartree–fock and local density-functional theories. *The Journal of Chemical Physics*, 98(2):1372–1377, 1993.
- [19] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–794, 2008.
- [20] Björn O Roos, Peter R Taylor, and Per EM Si. A complete active space scf method (casscf) using a density matrix formulated super-ci approach. *Chemical Physics*, 48(2):157–173, 1980.
- [21] Kerstin Andersson, Per-Åke Malmqvist, and Björn O Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, 96(2):1218–1226, 1992.
- [22] Per EM Siegbahn, Jan Almlöf, Anders Heiberg, and Björn O Roos. The complete active space scf (casscf) method in a newton–raphson formulation with application to the hno molecule. *The Journal of Chemical Physics*, 74(4):2384–2396, 1981.
- [23] Thierry Leininger, Hermann Stoll, Hans-Joachim Werner, and Andreas Savin. Combining long-range configuration interaction with short-range density functionals. *Chemical Physics Letters*, 275(3-4):151–160, 1997.
- [24] Jiali Gao, Adam Grofe, Haisheng Ren, and Peng Bao. Beyond kohn–sham approximation: Hybrid multistate wave function and density functional theory.

The Journal of Physical Chemistry Letters, 7(24):5143–5149, 2016.

- [25] Giovanni Li Manni, Rebecca K Carlson, Sijie Luo, Dongxia Ma, Jeppe Olsen, Donald G Truhlar, and Laura Gagliardi. Multiconfiguration pair-density functional theory. *Journal of Chemical Theory and Computation*, 10(9):3669–3680, 2014.
- [26] Yang Yang, Helen van Aggelen, and Weitao Yang. Double, rydberg and charge transfer excitations from pairing matrix fluctuation and particle-particle random phase approximation. *The Journal of Chemical Physics*, 139(22):224105, 2013.
- [27] Degao Peng, Helen van Aggelen, Yang Yang, and Weitao Yang. Linear-response time-dependent density-functional theory with pairing fields. *The Journal of Chemical Physics*, 140(18):18A522, 2014.
- [28] Yang Yang, Degao Peng, Jianfeng Lu, and Weitao Yang. Excitation energies from particle-particle random phase approximation: Davidson algorithm and benchmark studies. *The Journal of Chemical Physics*, 141(12):124104, 2014.
- [29] Yihan Shao, Martin Head-Gordon, and Anna I Krylov. The spin–flip approach within time-dependent density functional theory: Theory and applications to diradicals. *The Journal of Chemical Physics*, 118(11):4807–4818, 2003.
- [30] Fan Wang and Tom Ziegler. Time-dependent density functional theory based on a noncollinear formulation of the exchange-correlation potential. *The Journal of Chemical Physics*, 121(24):12191–12196, 2004.
- [31] Yves A Bernard, Yihan Shao, and Anna I Krylov. General formulation of spin-flip time-dependent density functional theory using non-collinear kernels: Theory, implementation, and benchmarks. *The Journal of Chemical Physics*, 136(20):204103, 2012.
- [32] Zhendong Li and Wenjian Liu. Theoretical and numerical assessments of spin-flip time-dependent density functional theory. *The Journal of Chemical Physics*, 136(2):024107, 2012.
- [33] Du Zhang, Degao Peng, Peng Zhang, and Weitao Yang. Analytic gradients, geometry optimization and excited state potential energy surfaces from the particle-particle random phase approximation. *Physical Chemistry Chemical Physics*, 17(2):1025–1038, 2015.
- [34] Ye Jin, Du Zhang, Zehua Chen, Neil Qiang Su, and Weitao Yang. Generalized optimized effective potential for orbital functionals: Self-consistent random

- phase approximations. Submitted, 2017.
- [35] Filipp Furche and Reinhart Ahlrichs. Adiabatic time-dependent density functional methods for excited state properties. *The Journal of Chemical Physics*, 117(16):7433–7447, 2002.
- [36] Qingxu Li, Qikai Li, and Zhigang Shuai. Local configuration interaction single excitation approach: Application to singlet and triplet excited states structure for conjugated chains. *Synthetic Metals*, 158(8):330–335, 2008.
- [37] Jie Liu and WanZhen Liang. Analytical hessian of electronic excited states in time-dependent density functional theory with tamm-dancoff approximation. *The Journal of Chemical Physics*, 135(1):014113, 2011.
- [38] An in-house program for quantum mechanical/molecular mechanical simulations used by dr. weitaoyang’s research group. <https://qm4d.org>.
- [39] George D Purvis III and Rodney J Bartlett. A full coupled-cluster singles and doubles model: the inclusion of disconnected triples. *The Journal of Chemical Physics*, 76(4):1910–1918, 1982.
- [40] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, B Mennucci, GA Petersson, et al. Gaussian 09 revision a. 1 gaussian inc; 2009. Wallingford CT.
- [41] Anna I Krylov. Size-consistent wave functions for bond-breaking: the equation-of-motion spin-flip model. *Chemical Physics Letters*, 338(4):375–384, 2001.
- [42] Anna I Krylov. Spin-flip configuration interaction: an electronic structure model that is both variational and size-consistent. *Chemical Physics Letters*, 350(5):522–530, 2001.
- [43] Anna I Krylov and C David Sherrill. Perturbative corrections to the equation-of-motion spin-flip self-consistent field model: Application to bond-breaking and equilibrium properties of diradicals. *The Journal of Chemical Physics*, 116(8):3194–3203, 2002.
- [44] John S Sears, C David Sherrill, and Anna I Krylov. A spin-complete version of the spin-flip approach to bond breaking: What is the impact of obtaining spin eigenfunctions? *The Journal of Chemical Physics*, 118(20):9084–9094, 2003.
- [45] David Casanova and Martin Head-Gordon. The spin-flip extended single excitation configuration interaction method. *The Journal of Chemical Physics*, 129(6):064104, 2008.

- [46] Wanyi Jiang, Chris C Jeffrey, and Angela K Wilson. Empirical correction of nondynamical correlation energy for density functionals. *The Journal of Physical Chemistry A*, 116(40):9969–9978, 2012.
- [47] Alexander Kramida and William C Martin. A compilation of energy levels and wavelengths for the spectrum of neutral beryllium (be i). *Journal of Physical and Chemical Reference Data*, 26(5):1185–1194, 1997.
- [48] CE Moore. Atomic energy levels; nbs circular 467/1; national bureau of standards: Washington, dc, 1949. *There is no corresponding record for this reference.*
- [49] Karol Kowalski and Piotr Piecuch. New coupled-cluster methods with singles, doubles, and noniterative triples for high accuracy calculations of excited electronic states. *The Journal of Chemical Physics*, 120(4):1715–1738, 2004.
- [50] A Kramida, Yu Ralchenko, and J Reader. Nist atomic spectra database (version 5.0). national institute of standards and technology, 2012.
- [51] Mark E Casida, Christine Jamorski, Kim C Casida, and Dennis R Salahub. 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. *The Journal of Chemical Physics*, 108(11):4439–4449, 1998.
- [52] Marko Schreiber, Mario R Silva-Junior, Stephan PA Sauer, and Walter Thiel. Benchmarks for electronically excited states: Caspt2, cc2, ccsd, and cc3. *The Journal of Chemical Physics*, 128(13):134110, 2008.
- [53] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136(3B):B864–B871, November 1964. doi: 10.1103/PhysRev.136.B864.
- [54] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A):A1133–A1138, November 1965. doi: 10.1103/PhysRev.140.A1133.
- [55] U. von Barth and L. Hedin. A local exchange-correlation potential for the spin polarized case. i. *J. Phys. C: Solid State Phys.*, 5(13):1629, 1972. ISSN 0022-3719. doi: 10.1088/0022-3719/5/13/012.
- [56] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can. J. Phys.*, 58(8):1200–1211, August 1980. ISSN 0008-4204. doi: 10.1139/p80-159.

- [57] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, September 1988. doi: 10.1103/PhysRevA.38.3098.
- [58] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, January 1988. doi: 10.1103/PhysRevB.37.785.
- [59] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45(23):13244–13249, June 1992. doi: 10.1103/PhysRevB.45.13244.
- [60] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868, October 1996. doi: 10.1103/PhysRevLett.77.3865.
- [61] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Localization and Delocalization Errors in Density Functional Theory and Implications for Band-Gap Prediction. *Phys. Rev. Lett.*, 100(14):146401, April 2008. doi: 10.1103/PhysRevLett.100.146401.
- [62] Erin R. Johnson, Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Delocalization errors in density functionals and implications for main-group thermochemistry. *J. Chem. Phys.*, 129(20):204112, November 2008. ISSN 0021-9606. doi: 10.1063/1.3021474.
- [63] Qin Wu and Weitao Yang. Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.*, 116(2):515–524, December 2001. ISSN 0021-9606. doi: 10.1063/1.1424928.
- [64] Axel D. Becke. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.*, 107(20):8554–8560, November 1997. ISSN 0021-9606. doi: 10.1063/1.475007.
- [65] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, April 1993. ISSN 0021-9606. doi: 10.1063/1.464913.
- [66] Axel D. Becke. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.*, 98(2):1372–1377, January 1993. ISSN 0021-9606. doi: 10.1063/1.464304.
- [67] Stefan Grimme. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.*, 124(3):034108, January 2006. ISSN

0021-9606. doi: 10.1063/1.2148954.

- [68] Sebastian Kozuch, David Gruzman, and Jan M. L. Martin. DSD-BLYP: A General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction. *J. Phys. Chem. C*, 114(48):20801–20808, December 2010. ISSN 1932-7447. doi: 10.1021/jp1070852.
- [69] Igor Ying Zhang, Neil Qiang Su, Éric A. G. Brémond, Carlo Adamo, and Xin Xu. Doubly hybrid density functional xDH-PBE0 from a parameter-free global hybrid model PBE0. *J. Chem. Phys.*, 136(17):174103, May 2012. ISSN 0021-9606. doi: 10.1063/1.3703893.
- [70] Neil Qiang Su and Xin Xu. Construction of a parameter-free doubly hybrid density functional from adiabatic connection. *J. Chem. Phys.*, 140(18):18A512, February 2014. ISSN 0021-9606. doi: 10.1063/1.4866457.
- [71] Narbe Mardirossian and Martin Head-Gordon. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.*, 115(19):2315–2372, October 2017. ISSN 0026-8976. doi: 10.1080/00268976.2017.1333644.
- [72] Oleg A. Vydrov and Troy Van Voorhis. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.*, 133(24):244103, December 2010. ISSN 0021-9606. doi: 10.1063/1.3521275.
- [73] Mark S. Gordon, Dmitri G. Fedorov, Spencer R. Pruitt, and Lyudmila V. Slipchenko. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.*, 112(1):632–672, January 2012. ISSN 0009-2665. doi: 10.1021/cr200093j.
- [74] Ryan M. Richard, Ka Un Lao, and John M. Herbert. Aiming for Benchmark Accuracy with the Many-Body Expansion. *Acc. Chem. Res.*, 47(9):2828–2836, September 2014. ISSN 0001-4842. doi: 10.1021/ar500119q.
- [75] Michael A. Collins and Ryan P. A. Bettens. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.*, 115(12):5607–5642, June 2015. ISSN 0009-2665. doi: 10.1021/cr500455b.
- [76] Hermann Stoll. Correlation energy of diamond. *Phys. Rev. B*, 46(11):6700–6704, September 1992. doi: 10.1103/PhysRevB.46.6700.
- [77] Beate Paulus. The method of increments—a wavefunction-based ab initio correlation method for solids. *Physics Reports*, 428(1):1–52, May 2006. ISSN 0370-1573. doi: 10.1016/j.physrep.2006.01.003.

- [78] F. L. Hirshfeld. XVII. Spatial Partitioning of Charge Density. *Isr. J. Chem.*, 16(2-3):198–201, 1977. ISSN 1869-5868. doi: 10.1002/ijch.197700033.
- [79] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132(15):154104, April 2010. ISSN 0021-9606. doi: 10.1063/1.3382344.
- [80] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, May 2011. ISSN 0192-8651. doi: 10.1002/jcc.21759.
- [81] Alexandre Tkatchenko, Robert A. DiStasio, Roberto Car, and Matthias Scheffler. Accurate and Efficient Method for Many-Body van der Waals Interactions. *Phys. Rev. Lett.*, 108(23):236402, June 2012. doi: 10.1103/PhysRevLett.108.236402.
- [82] Qin Wu and Weitao Yang. A direct optimization method for calculating density functionals and exchange–correlation potentials from electron densities. *J. Chem. Phys.*, 118(6):2498–2509, January 2003. ISSN 0021-9606. doi: 10.1063/1.1535422.
- [83] Jörg Behler and Michele Parrinello. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.*, 98(14):146401, April 2007. doi: 10.1103/PhysRevLett.98.146401.
- [84] Matthew Welborn, Lixue Cheng, and Thomas F. Miller. Transferability in Machine Learning for Electronic Structure via the Molecular Orbital Basis. *J. Chem. Theory Comput.*, 14(9):4772–4779, September 2018. ISSN 1549-9618. doi: 10.1021/acs.jctc.8b00636.
- [85] J. P. Coe. Machine Learning Configuration Interaction. *J. Chem. Theory Comput.*, October 2018. ISSN 1549-9618. doi: 10.1021/acs.jctc.8b00849.
- [86] LiHong Hu, XiuJun Wang, LaiHo Wong, and GuanHua Chen. Combined first-principles calculation and neural-network correction approach for heat of formation. *J. Chem. Phys.*, 119(22):11501–11507, November 2003. ISSN 0021-9606. doi: 10.1063/1.1630951.
- [87] John C. Snyder, Matthias Rupp, Katja Hansen, Klaus-Robert Müller, and Kieron Burke. Finding Density Functionals with Machine Learning. *Phys. Rev. Lett.*, 108(25):253002, June 2012. doi: 10.1103/PhysRevLett.108.253002.

- [88] Yan Zhao and Donald G. Truhlar. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and non-covalent interactions. *J. Chem. Phys.*, 125(19):194101, November 2006. ISSN 0021-9606. doi: 10.1063/1.2370993.
- [89] Yan Zhao and Donald G. Truhlar. 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 Account*, 120(1):215–241, May 2008. ISSN 1432-2234. doi: 10.1007/s00214-007-0310-x.
- [90] Narbe Mardirossian and Martin Head-Gordon. ω B97X-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.*, 16(21):9904–9924, 2014. doi: 10.1039/C3CP54374A.
- [91] Narbe Mardirossian and Martin Head-Gordon. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.*, 144(21):214110, June 2016. ISSN 0021-9606. doi: 10.1063/1.4952647.
- [92] Albert P. Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Phys. Rev. B*, 87(18):184115, May 2013. doi: 10.1103/PhysRevB.87.184115.
- [93] Lars Goerigk, Andreas Hansen, Christoph Bauer, Stephan Ehrlich, Asim Najibi, and Stefan Grimme. 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.*, 19(48):32184–32215, 2017. doi: 10.1039/C7CP04913G.
- [94] George D. Purvis and Rodney J. Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76(4): 1910–1918, February 1982. ISSN 0021-9606. doi: 10.1063/1.443164.
- [95] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd,

- E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 Revision A.03, 2016. Gaussian Inc. Wallingford CT.
- [96] Dmitriy Rappoport and Filipp Furche. Property-optimized gaussian basis sets for molecular response calculations. *The Journal of chemical physics*, 133(13):134105, 2010.
- [97] Chen Li, Xiao Zheng, Neil Qiang Su, and Weitao Yang. Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations. *Natl Sci Rev*, 5(2):203–215, March 2018. ISSN 2095-5138. doi: 10.1093/nsr/nwx111.
- [98] Neil Qiang Su, Chen Li, and Weitao Yang. Describing strong correlation with fractional-spin correction in density functional theory. *PNAS*, 115(39):9678–9683, September 2018. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1807095115.
- [99] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136(3B):B864–B871, November 1964. doi: 10.1103/PhysRev.136.B864.
- [100] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A):A1133–A1138, November 1965. doi: 10.1103/PhysRev.140.A1133.
- [101] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry. Oxford University Press, Oxford, New York, May 1994. ISBN 978-0-19-509276-9.
- [102] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can. J. Phys.*, 58(8):1200–1211, August 1980. ISSN 0008-4204. doi: 10.1139/p80-159.
- [103] John P. Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45(23):13244–13249, June 1992. doi: 10.1103/PhysRevB.45.13244.
- [104] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, September 1988. doi: 10.1103/PhysRevA.38.3098.

- [105] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, January 1988. doi: 10.1103/PhysRevB.37.785.
- [106] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868, October 1996. doi: 10.1103/PhysRevLett.77.3865.
- [107] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.*, 98(45):11623–11627, November 1994. ISSN 0022-3654, 1541-5740. doi: 10.1021/j100096a001.
- [108] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, 110(13):6158–6170, March 1999. ISSN 0021-9606. doi: 10.1063/1.478522.
- [109] Matthias Ernzerhof and Gustavo E. Scuseria. Assessment of the Perdew–Burke–Ernzerhof exchange–correlation functional. *J. Chem. Phys.*, 110(11):5029–5036, March 1999. ISSN 0021-9606. doi: 10.1063/1.478401.
- [110] Chen Li, Xiao Zheng, Neil Qiang Su, and Weitao Yang. Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations. *Natl. Sci. Rev.*, 5(2):203–215, March 2018. ISSN 2095-5138, 2053-714X. doi: 10.1093/nsr/nwx111.
- [111] John P. Perdew, Robert G. Parr, Mel Levy, and Jose L. Balduz. Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy. *Phys. Rev. Lett.*, 49(23):1691–1694, December 1982. doi: 10.1103/PhysRevLett.49.1691.
- [112] Weitao Yang, Yingkai Zhang, and Paul W. Ayers. Degenerate Ground States and a Fractional Number of Electrons in Density and Reduced Density Matrix Functional Theory. *Phys. Rev. Lett.*, 84(22):5172–5175, May 2000. doi: 10.1103/PhysRevLett.84.5172.
- [113] Yingkai Zhang and Weitao Yang. Perspective on “Density-functional theory for fractional particle number: Derivative discontinuities of the energy”. In Christopher J. Cramer and Donald G. Truhlar, editors, *Theoretical Chemistry Accounts: New Century Issue*, pages 346–348. Springer, Berlin, Heidelberg, 2001. ISBN 978-3-662-10421-7.
- [114] Weitao Yang, Aron J. Cohen, and Paula Mori-Sánchez. Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional

- theory. *J. Chem. Phys.*, 136(20):204111, May 2012. ISSN 0021-9606. doi: 10.1063/1.3702391.
- [115] Anthony D. Dutoi and Martin Head-Gordon. Self-interaction error of local density functionals for alkali-halide dissociation. *Chem. Phys. Lett.*, 422(1): 230–233, April 2006. ISSN 0009-2614. doi: 10.1016/j.cplett.2006.02.025.
- [116] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.*, 125(20): 201102, November 2006. ISSN 0021-9606. doi: 10.1063/1.2403848.
- [117] Adrienn Ruzsinszky, John P. Perdew, Gábor I. Csonka, Oleg A. Vydrov, and Gustavo E. Scuseria. Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals. *J. Chem. Phys.*, 125(19):194112, November 2006. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.2387954.
- [118] Oleg A. Vydrov, Gustavo E. Scuseria, and John P. Perdew. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.*, 126(15):154109, April 2007. ISSN 0021-9606. doi: 10.1063/1.2723119.
- [119] Xiao Zheng, Min Liu, Erin R. Johnson, Julia Contreras-García, and Weitao Yang. Delocalization error of density-functional approximations: A distinct manifestation in hydrogen molecular chains. *J. Chem. Phys.*, 137(21):214106, December 2012. ISSN 0021-9606. doi: 10.1063/1.4768673.
- [120] F. Flores, J. Ortega, and H. Vázquez. Modelling energy level alignment at organic interfaces and density functional theory. *Phys. Chem. Chem. Phys.*, 11(39):8658–8675, 2009. doi: 10/bcd3w4. URL <https://pubs.rsc.org/en/content/articlelanding/2009/cp/b902492c>.
- [121] A. M. Souza, I. Rungger, C. D. Pemmaraju, U. Schwingenschloegl, and S. Sanvito. Constrained-DFT method for accurate energy-level alignment of metal/molecule interfaces. *Phys. Rev. B*, 88(16):165112, October 2013. doi: 10/gfrh7c. URL <https://link.aps.org/doi/10.1103/PhysRevB.88.165112>.
- [122] Gianfranco Pacchioni. First Principles Calculations on Oxide-Based Heterogeneous Catalysts and Photocatalysts: Problems and Advances. *Catal. Lett.*, 145(1):80–94, January 2015. ISSN 1011-372X, 1572-879X. doi: 10/gfrh73. URL <https://link.springer.com/article/10.1007/s10562-014-1386-2>.
- [123] David A. Egger, Zhen-Fei Liu, Jeffrey B. Neaton, and Leeor Kronik. Reliable Energy Level Alignment at Physisorbed Molecule-Metal Interfaces from Density Functional Theory. *Nano Lett.*, 15(4):2448–2455, April 2015. ISSN

1530-6984. doi: 10/f3sv86. URL <https://doi.org/10.1021/nl504863r>.

- [124] Shenghao Wang, Takeaki Sakurai, Weijia Wen, and Yabing Qi. Energy Level Alignment at Interfaces in Metal Halide Perovskite Solar Cells. *Adv. Mater. Interfaces*, 5(22):1800260, 2018. ISSN 2196-7350. doi: 10/gdxs3x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/admi.201800260>.
- [125] Franziska Simone Hegner, Drialys Cardenas-Morcoso, Sixto Giménez, Núria López, and Jose Ramon Galan-Mascaros. Level Alignment as Descriptor for Semiconductor/Catalyst Systems in Water Splitting: The Case of Hematite/Cobalt Hexacyanoferrate Photoanodes. *ChemSusChem*, 10(22):4552–4560, 2017. ISSN 1864-564X. doi: 10/gbzrc3. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cssc.201701538>.
- [126] A. Savin and Hj Flad. Density Functionals for the Yukawa Electron-Electron Interaction. *Int. J. Quantum Chem.*, 56(4):327–332, November 1995. ISSN 0020-7608. doi: 10.1002/qua.560560417.
- [127] Peter M. W. Gill, Ross D. Adamson, and John A. Pople. Coulomb-attenuated exchange energy density functionals. *Mol. Phys.*, 88(4):1005–1009, July 1996. ISSN 0026-8976. doi: 10.1080/00268979609484488.
- [128] Thierry Leininger, Hermann Stoll, Hans-Joachim Werner, and Andreas Savin. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275(3-4):151–160, August 1997. ISSN 00092614. doi: 10.1016/S0009-2614(97)00758-6.
- [129] Hisayoshi Iikura, Takao Tsuneda, Takeshi Yanai, and Kimihiko Hirao. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.*, 115(8):3540–3544, August 2001. ISSN 0021-9606. doi: 10.1063/1.1383587.
- [130] J. Toulouse, F. Colonna, and A. Savin. Long-range-short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A*, 70(6):062505, December 2004. ISSN 2469-9926. doi: 10.1103/PhysRevA.70.062505.
- [131] Takeshi Yanai, David P Tew, and Nicholas C Handy. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.*, 393(1):51–57, July 2004. ISSN 0009-2614. doi: 10.1016/j.cplett.2004.06.011.
- [132] Roi Baer and Daniel Neuhauser. Density Functional Theory with Correct Long-Range Asymptotic Behavior. *Phys. Rev. Lett.*, 94(4):043002, February 2005.

doi: 10.1103/PhysRevLett.94.043002.

- [133] Oleg A. Vydrov and Gustavo E. Scuseria. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.*, 125(23):234109, December 2006. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.2409292.
- [134] Oleg A. Vydrov, Jochen Heyd, Aliaksandr V. Krukau, and Gustavo E. Scuseria. Importance of short-range versus long-range Hartree-Fock exchange for the performance of hybrid density functionals. *J. Chem. Phys.*, 125(7):074106, August 2006. ISSN 0021-9606. doi: 10.1063/1.2244560.
- [135] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Development of exchange-correlation functionals with minimal many-electron self-interaction error. *J. Chem. Phys.*, 126(19):191109, May 2007. ISSN 0021-9606. doi: 10.1063/1.2741248.
- [136] Jeng-Da Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.*, 128(8):084106, February 2008. ISSN 0021-9606. doi: 10.1063/1.2834918.
- [137] Roi Baer, Ester Livshits, and Ulrike Salzner. Tuned Range-Separated Hybrids in Density Functional Theory. *Annu. Rev. Phys. Chem.*, 61(1):85–109, 2010. doi: 10.1146/annurev.physchem.012809.103321.
- [138] Yan Zhao, Benjamin J. Lynch, and Donald G. Truhlar. Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A*, 108(21):4786–4791, May 2004. ISSN 1089-5639. doi: 10.1021/jp049253v.
- [139] Stefan Grimme. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.*, 124(3):034108, January 2006. ISSN 0021-9606. doi: 10.1063/1.2148954.
- [140] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.*, 131(17):174105, November 2009. ISSN 0021-9606. doi: 10.1063/1.3244209.
- [141] Ying Zhang, Xin Xu, and William A. Goddard. Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics. *PNAS*, 106(13):4963–4968, March 2009. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0901093106.
- [142] Neil Qiang Su, Weitao Yang, Paula Mori-Sánchez, and Xin Xu. Fractional Charge Behavior and Band Gap Predictions with the XYG3 Type of Doubly

- Hybrid Density Functionals. *J. Phys. Chem. A*, 118(39):9201–9211, October 2014. ISSN 1089-5639. doi: 10.1021/jp5029992.
- [143] Neil Qiang Su, Aaron Mahler, and Weitao Yang. Preserving Symmetry and Degeneracy in the Localized Orbital Scaling Correction Approach. *J. Phys. Chem. Lett.*, 11(4):1528–1535, February 2020. doi: 10.1021/acs.jpcclett.9b03888.
- [144] Yuncai Mei, Chen Li, Neil Qiang Su, and Weitao Yang. Approximating Quasiparticle and Excitation Energies from Ground State Generalized Kohn–Sham Calculations. *J. Phys. Chem. A*, 123(3):666–673, January 2019. ISSN 1089-5639, 1520-5215. doi: 10.1021/acs.jpca.8b10380.
- [145] Yuncai Mei and Weitao Yang. Excited-State Potential Energy Surfaces, Conical Intersections, and Analytical Gradients from Ground-State Density Functional Theory. *J. Phys. Chem. Lett.*, 10(10):2538–2545, May 2019. ISSN 1948-7185. doi: 10.1021/acs.jpcclett.9b00712.
- [146] Yingkai Zhang and Weitao Yang. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. *J. Chem. Phys.*, 109(7):2604–2608, August 1998. ISSN 0021-9606. doi: 10.1063/1.476859.
- [147] Xiao Zheng, Aron J. Cohen, Paula Mori-Sánchez, Xiangqian Hu, and Weitao Yang. Improving Band Gap Prediction in Density Functional Theory from Molecules to Solids. *Phys. Rev. Lett.*, 107(2):026403, July 2011. doi: 10.1103/PhysRevLett.107.026403.
- [148] Mark S. Gordon and Michael W. Schmidt. Chapter 41 - Advances in electronic structure theory: GAMESS a decade later. In Clifford E. Dykstra, Gernot Frenking, Kwang S. Kim, and Gustavo E. Scuseria, editors, *Theory and Applications of Computational Chemistry*, pages 1167–1189. Elsevier, Amsterdam, January 2005. ISBN 978-0-444-51719-7. doi: 10.1016/B978-044451719-7/50084-6.
- [149] Michael W. Schmidt, Kim K. Baldridge, Jerry A. Boatz, Steven T. Elbert, Mark S. Gordon, Jan H. Jensen, Shiro Koseki, Nikita Matsunaga, Kiet A. Nguyen, Shujun Su, Theresa L. Windus, Michel Dupuis, and John A. Montgomery. General atomic and molecular electronic structure system. *J. Comput. Chem.*, 14(11):1347–1363, 1993. ISSN 1096-987X. doi: 10.1002/jcc.540141112.
- [150] Volker Blum, Ralf Gehrke, Felix Hanke, Paula Havu, Ville Havu, Xinguo Ren, Karsten Reuter, and Matthias Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.*, 180(11):2175–2196, November 2009. ISSN 0010-4655. doi: 10.1016/j.cpc.2009.06.022.

- [151] Xinguo Ren, Patrick Rinke, Volker Blum, Jürgen Wieferink, Alexandre Tkatchenko, Andrea Sanfilippo, Karsten Reuter, and Matthias Scheffler. Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys.*, 14(5):053020, May 2012. ISSN 1367-2630. doi: 10.1088/1367-2630/14/5/053020.
- [152] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, April 1993. ISSN 0021-9606. doi: 10.1063/1.464913.
- [153] Peter J. Knowles and Hans-Joachim Werner. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chemical Physics Letters*, 145(6):514–522, April 1988. ISSN 00092614. doi: 10.1016/0009-2614(88)87412-8.
- [154] Stephen R. Langhoff and Ernest R. Davidson. Configuration interaction calculations on the nitrogen molecule. *Int. J. Quantum Chem.*, 8(1):61–72, 1974. ISSN 1097-461X. doi: 10.1002/qua.560080106.
- [155] Hans-Joachim Werner and Peter J. Knowles. An efficient internally contracted multiconfiguration–reference configuration interaction method. *J. Chem. Phys.*, 89(9):5803–5814, November 1988. ISSN 0021-9606. doi: 10.1063/1.455556.
- [156] R. S. Mulliken. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J. Chem. Phys.*, 23(10):1833–1840, October 1955. ISSN 0021-9606. doi: 10.1063/1.1740588.
- [157] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, 32(7):1456–1465, 2011. ISSN 1096-987X. doi: 10.1002/jcc.21759.
- [158] Chr. Møller and M. S. Plesset. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.*, 46(7):618–622, October 1934. doi: 10.1103/PhysRev.46.618.
- [159] D. G. Streets, W. Elane Hall, and Gerald P. Ceasar. Mesomeric mixing in the π energy levels of amino-benzenes studied by photoelectron spectroscopy. *Chem. Phys. Lett.*, 17(1):90–94, November 1972. ISSN 0009-2614. doi: 10.1016/0009-2614(72)80334-8.
- [160] Dmitry Khuseynov, Matthew T. Fontana, and Andrei Sanov. Photoelectron spectroscopy and photochemistry of tetracyanoethylene radical anion in the gas phase. *Chem. Phys. Lett.*, 550:15–18, October 2012. ISSN 0009-2614. doi:

10.1016/j.cplett.2012.08.035.

- [161] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Fractional charge perspective on the band gap in density-functional theory. *Phys. Rev. B*, 77(11), March 2008. ISSN 1098-0121, 1550-235X. doi: 10.1103/PhysRevB.77.115123. URL <https://link.aps.org/doi/10.1103/PhysRevB.77.115123>.

Biography

Zehua Chen received his B.S. degree in Chemistry at Peking University in 2015, majoring in theoretical chemistry. He later received the Ph.D. degree in theoretical chemistry at Duke University in 2021. In 2019, Zehua received support from the Kathleen Zielik fellowship. Zehua has given two oral presentations in ACS national meetings, which include:

- August 2019: Development of a finite-range non-local functional, assisted by machine-learning.
- March 2018: Multireference density functional theory with generalized auxiliary reference systems for ground and excited states.

He also gave a poster on the same topic of the 2018 ACS talk in 2017 during the ACTC meeting. Below is a list of publications by Zehua Chen:

- Mei, Y., **Chen, Z.**, & Yang, W. Self-Consistent Calculation of the Localized Orbital Scaling Correction for Correct Electron Densities and Energy-Level Alignments in Density Functional Theory. *The Journal of Physical Chemistry Letters*, **2020**, *11*(23), 10269-10277.
- Jin, Y., Su, N. Q., **Chen, Z.**, & Yang, W. Introductory lecture: when the density of the noninteracting reference system is not the density of the physical system in density functional theory. *Faraday Discussions*, **2020**, *224*, 9-26.
- Li, G., **Chen, Z.**, Li, Y., Zhang, D., Yang, W., Liu, Y., & Cao, L. Engineering substrate interaction to improve hydrogen evolution catalysis of monolayer MoS₂ films beyond Pt. *ACS nano*, **2020**, *14*(2), 1707-1714.

- Pinter, B., Al-Saadon, R., Chen, Z., & Yang, W. Spin-state energetics of iron (II) porphyrin from the particle-particle random phase approximation. *The European Physical Journal B*, **2018**, 91(11), 1-10.
- Jin, Y., Zhang, D., Chen, Z., Su, N. Q., & Yang, W. Generalized optimized effective potential for orbital functionals and self-consistent calculation of random phase approximations. *The Journal of Physical Chemistry Letters*, **2017**, 8(19), 4746-4751.
- Chen, Z., Zhang, D., Jin, Y., Yang, Y., Su, N. Q., & Yang, W. Multireference density functional theory with generalized auxiliary systems for ground and excited states. *The Journal of Physical Chemistry Letters*, **2017**, 8(18), 4479-4485.
- Wang, Y. C., Chen, Z. H., & Jiang, H. The local projection in the density functional theory plus U approach: a critical assessment. *The Journal of Chemical Physics*, **2016**, 144(14), 144106.