Systematic Improvement and Applications of Density Functional Theory in Complex Chemical Systems

by

Tianyu Zhu

B.S., University of Science and Technology of China (2013)

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

Signature redacted

Author .....
Department of Chemistry
May 3, 2018

Signature redacted

Certified by...
Troy Van Voorhis
Haslam and Dewey Professor of Chemistry
Thesis Supervisor

Signature redacted

Accepted by .....
Robert W. Field
Chairman, Department Committee on Graduate Theses

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Signature redacted

Professor Jianshu Cao ....
Chairman, Thesis Committee
Professor of Chemistry

Signature redacted

Professor Troy Van Voorhis ....
Thesis Supervisor
Haslam and Dewey Professor of Chemistry

Signature redacted

Professor Adam P. Willard. .....
Member, Thesis Committee
Assistant Professor of Chemistry

Systematic Improvement and Applications of Density Functional Theory in Complex Chemical Systems

by

Tianyu Zhu

Submitted to the Department of Chemistry
on May 3, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemistry

Abstract

Density functional theory (DFT) has become the *de facto* method for quantum mechanical simulations of molecules and solids because of its high performance to cost ratio. In this thesis, we discuss two aspects of DFT simulations in complex chemical systems: systematic improvement of the accuracy of density functional approximations and proper utilization of DFT methods for efficient modeling of electronic properties.

We first develop the many-pair expansion (MPE) method, which is a density functional hierarchy that systematically corrects any deficiencies of an approximate density functional to converge to the exact energy. We show that MPE gives accurate results for 1D/2D Hubbard and 1D Peierls-Hubbard models, suggesting its ability to remove strong correlation errors. Applying MPE to unsaturated hydrocarbons in the Pariser-Parr-Pople lattice model, we find that it deals very well with dispersion interactions. Afterwards, we describe our efforts to implement MPE for molecular systems. A new density decomposition method, self-attractive Hartree (SAH), is developed to generate localized and smooth fragment densities. The SAH decomposition is shown to be useful for extracting chemical bonding information directly from the electron density and further applied to develop a simple and accurate hydrogen bonding strength indicator. Using SAH fragment densities, we demonstrate that MPE provides accurate description of reaction energies and bond breaking processes for a few small molecules, even with a low-level starting functional and low orders of expansion.

To show how DFT methods can be properly utilized to obtain electronic properties of interest, we employ the theoretical investigation of organic light-emitting diodes (OLEDs) as an example. We adopt a hybrid quantum mechanics/molecular mechanics (QM/MM) approach to reveal the charge and energy mechanisms of a host-guest phosphorescent OLED in condensed phase, emphasizing the importance of incorporating environment effects. We then show successful computational design of new thermally activated delayed fluorescence (TADF) materials using conventional time-dependent DFT method, while point out the need for better excited-state DFT methods. Finally, we develop efficient computational screening protocols to study TADF materials based on a restricted open-shell Kohn-Sham approach.

Thesis Supervisor: Troy Van Voorhis
Title: Haslam and Dewey Professor of Chemistry

Acknowledgments

First and foremost, I would like to thank my advisor Prof. Troy Van Voorhis, for his continuous support, encouragement, and insightful advice throughout my Ph.D. I am grateful to Troy for introducing me to the world of electronic structure with his immense knowledge and always guiding me with great patience in my research work. I also thank Troy for offering me so many opportunities and much freedom to work on various scientific topics and learn from different people. It has been a great honor to be his student.

I would also like to acknowledge my thesis committee members, Prof. Jianshu Cao and Prof. Adam Willard, for their time, encouragement and valuable comments on my research.

The members of the Van Voorhis group have provided me with such a friendly and helpful study and work environment. Particularly, my thanks go to Piotr de Silva, who was always willing to offer valuable suggestions on my research and career. It was a great pleasure to collaborate with him and I learned a lot from the discussions with him. I want to thank Shuhao Wen, David McMahon, Matthew Welborn and Michael Mavros, for their help and guidance when I started my graduate research. I also appreciate all the collaborations, scientific discussions, and kind help from Zhou Lin, Hongzhou Ye and Diptarka Hait. It is a very enjoyable experience working with all other past and present members, and my thanks go to Takashi Tsuchimochi, Eric Hontz, Helen Xie, Valerie Vaissier Welborn, James Shepherd, Helen van Aggelen, Pedro Henrique de Oliveira Neto, Alex Kohn, Nadav Geva, Tamar Goldzack, Nathan Ricke, Lexie McIsaac, Ricardo Pedro, Andrew Kim, Henry Tran and Hikari Iwasaki. I also want to acknowledge Li Miao for her administrative help.

I am appreciative of the collaborations with many excellent experimentalists. I want to express my gratitudes to Chi Zhang, Peng Dai, Prof. Bradley Pentelute, Wenliang Huang, Markus Einzinger, Prof. Stephen Buchwald, Prof. Marc Baldo, Katsuaki Kawasumi, Tony Wu and Prof. Timothy Swager. I have learned so much not only from their fascinating scientific ideas, but also from their enthusiasms and approaches for conducting research.

Last but by no means least, I am very much indebted to my parents for their love, understanding and supporting me in all my pursuits. I am deeply grateful to my girlfriend, Yu Fang, for encouraging me through all tough times and all sacrifices she has made. Her love and support has always been my great strength. Without them, I could not have completed my Ph.D. This thesis is dedicated to them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The application of quantum mechanics to the hydrogen molecule by Heitler and London in 1927 marks the first quantum chemical calculation of the molecule [1]. The primary goal of quantum chemistry is to solve the motions of electrons in atoms and molecules. During more than 90 years of development, the main efforts in this field have been made to search for faster and more accurate approximate solutions to the electronic Schrödinger equation [2, 3]. Nowadays, quantum chemistry has become a significant part of chemical and materials science research. It not only provides understanding and explanations of material properties and reaction mechanisms, but also begins to give accurate predictions and guidance to experiments.

Quantum chemistry methods are especially suitable for calculating the geometrical structures [4], thermochemistry [5], reaction barriers [6], spectroscopy [7] and other electronic properties of molecular and solid materials. The advancement in algorithms and computing power makes it possible to accurately treat systems containing hundreds or even thousands of atoms [8, 9]. The quantum chemistry tools are thus applied to study important and complex chemical systems in a wide range of applications [10]. Some of these applications are the understanding and design of catalytic processes in materials and enzymes [11, 12, 13], solar energy harvesting and conversions [14, 15], the design of electroluminescent emitters [16, 17] and the understanding of active-site chemistry [18, 19, 20].

Without the use of density functional theory (DFT) [21, 22], many of these applications would be intractable. Although DFT may not provide the benchmark-level accuracy as some other quantum chemistry methods, its high computational accuracy to cost ratio

makes it the workhorse for quantum chemical simulations. Despite much success in its formulations and applications, DFT still suffers from notable systematic errors due to the use of approximate density functionals [23]. For example, common density functionals cannot accurately describe the bond breaking processes in molecules. In reality, one encounters an ever-growing list of specific density functionals that have been tailored to particular physical conditions [24, 25, 26]. That is to say, there is no known universal density functional that can deal with all chemical applications and no method exists to systematically improve an approximate density functional. This is highly undesired from a practical perspective, as it would be very difficult for an inexperienced DFT user to choose the right functional for a given application. In addition, even with the appropriate functional, DFT simulations may still be insufficient to accurately predict electronic properties in complex chemical systems. For instance, to understand the charge transfer mechanisms in disordered organic semiconductors, one needs methods to correctly incorporate the condensed phase effects and accurately estimate the excited-state energies of materials.

In this thesis, we would like to tackle the above problems from two aspects. First, we develop a new DFT-based approach, many-pair expansion (MPE), that is capable of systematically correcting deficiencies in any approximate density functional and finally converges to the exact energy. The resulting method is shown to improve the performance of approximate DFT calculations in a series of important chemical problems. Second, we show how DFT methods can be used properly for efficient condensed phase modeling of organic light-emitting diodes (OLEDs) [27] and computational screening of thermally activated delayed fluorescence (TADF) materials [28]. The resulting simulations lead to understanding of important charge and energy transfer mechanisms in OLED devices and design of new highly efficient TADF emitters.

In this chapter, we review some of the underlying quantum chemistry and condensed phase simulation tools for this work. We start from the wavefunction theory including Hartree-Fock and configuration interaction methods, and then extensively introduce the Kohn-Sham (KS) density functional theory. Except for regular KS-DFT methods, we also review several unconventional DFT methods that are useful for our method development and OLED applications. We then turn our focus to incorporating environment effects in DFT simulations, where we review the implicit and explicit environment models for efficient condensed phased modeling of complex chemical systems. Finally, we outline the structure

for the rest of this thesis.

1.1 Wavefunction theory

The main interest of quantum chemistry methods is to find solutions to non-relativistic time-independent Schrödinger equation:

$$\hat{H}\Psi(\{\mathbf{r_i}\},\{\mathbf{R_A}\}) = E\Psi(\{\mathbf{r_i}\},\{\mathbf{R_A}\}) \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian of the system, $\Psi(\{\mathbf{r_i}\},\{\mathbf{R_A}\})$ is the wavefunction of the system, E is the total energy, $\{\mathbf{r_i}\}$ and $\{\mathbf{R_A}\}$ represent the positions of electrons and nuclei in the system. The Hamiltonian $\hat{H}$ can be expressed as:

$$\hat{H} = -\frac{\hbar^2}{2M_A}\sum_A \nabla_A^2 - \frac{\hbar^2}{2m}\sum_i \nabla_i^2 + \sum_{A<B}\frac{Z_A Z_B e^2}{4\pi\epsilon_0 R_{AB}} - \sum_{A,i}\frac{Z_A e^2}{4\pi\epsilon_0 r_{Ai}} + \sum_{i<j}\frac{e^2}{4\pi\epsilon_0 r_{ij}}, \tag{1.2}$$

where i, j refer to electrons and A, B refer to nuclei. For simplicity, we adopt the atomic units from now on, so the Hamiltonian can be simplified as:

$$\hat{H} = -\frac{1}{2M_A}\sum_A \nabla_A^2 - \frac{1}{2}\sum_i \nabla_i^2 + \sum_{A<B}\frac{Z_A Z_B}{R_{AB}} - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{i<j}\frac{1}{r_{ij}}. \tag{1.3}$$

The five terms in Eq. 1.3 are the kinetic energy operator for nuclei and electrons, repulsion between nuclei, attraction between nuclei and electrons, repulsion between electrons, respectively.

Since nuclei are much heavier than electrons, electrons move much faster than nuclei. One can then consider the electrons to be moving in the field of fixed nuclei in a molecule. This leads to Born-Oppenheimer (BO) approximation [29], which assumes the motions of nuclei and electrons can be separated. One can first freeze the nuclei positions so that the kinetic energy of nuclei is neglected and the attraction between nuclei becomes a constant in Eq. 1.3. Such assumption results in the electronic Schrödinger equation:

$$\hat{H}_{el}\Psi(\{\mathbf{r_i}\};\{\mathbf{R_A}\}) = E_{el}(\{\mathbf{R_A}\})\Psi(\{\mathbf{r_i}\};\{\mathbf{R_A}\}), \tag{1.4}$$

with the electronic Hamiltonian

$$\hat{H}_{el} = -\frac{1}{2}\sum_{i}\nabla_i^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{i<j}\frac{1}{r_{ij}}. \tag{1.5}$$

Solving the electronic Schrödinger equation yields the electronic energy $E_{el}(\{\mathbf{R_A}\})$ and electronic wavefunction $\Psi(\{\mathbf{r_i}\};\{\mathbf{R_A}\})$, both of which depend parametrically on the nuclear coordinates $\{\mathbf{R_A}\}$. The total energy of the system with fixed nuclear positions $\{\mathbf{R_A}\}$ can thus be computed as:

$$E_{tot}(\{\mathbf{R_A}\}) = E_{el}(\{\mathbf{R_A}\}) + \sum_{A<B}\frac{Z_A Z_B}{R_{AB}}. \tag{1.6}$$

To solve the full Schrödinger equation in Eq. 1.1, one can construct the nuclear Schrödinger equation to describe the motion of nuclei:

$$\hat{H}_{nuc}\chi(\{\mathbf{R_A}\}) = E\chi(\{\mathbf{R_A}\}) \tag{1.7}$$

with the nuclear Hamiltonian

$$\hat{H}_{nuc} = -\frac{1}{2M_A}\sum_{A}\nabla_A^2 + E_{el}(\{\mathbf{R_A}\}) + \sum_{A<B}\frac{Z_A Z_B}{R_{AB}}. \tag{1.8}$$

Note that the electronic total energy $E_{tot}(\{\mathbf{R_A}\})$ in Eq. 1.6 provides the potential for the nuclear motions. The solution to Eq. 1.7 gives the ground-state energy of the whole system and the nuclear wavefunction $\chi(\{\mathbf{R_A}\})$.

Specially, the focus of quantum chemistry methods is to solve the electronic Schrödinger equation described in Eqs. 1.4-1.6, which is also central to the methods and applications in this thesis. Therefore, we will only further review and discuss the efforts for solving the electronic Schrödinger equation from now on.

1.1.1 Hartree-Fock method

The Schrödinger equation provides an exact way to describe a quantum system, but it is too complicated to be solved. As commented by Paul Dirac [30], "the underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application

of these laws leads to equations much too complicated to be soluble." In reality, one always has to make approximations to solve the electronic Schrödinger equation efficiently. Among all such approximations, Hartree-Fock (HF) is the most simple one but also provides the foundation for other more complicated quantum chemistry methods. In this section, we review the formalism of the Hartree-Fock method and we follow the derivation in Ref. [3].

Before we delve into the Hartree-Fock method, we need to find a way to represent the many-electron electronic wavefunction $\Psi(\{\mathbf{r_i}\};\{\mathbf{R_A}\})$. First, we add a spin coordinate ω for electron spins into the wavefunction, and two spin functions $\alpha(\omega)$ (spin up) or $\beta(\omega)$ (spin down) are introduced. The new coordinate is $\mathbf{x} = \{\mathbf{r}, \omega\}$ and the electronic wavefunction is $\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$ for an N-electron system. Note that we drop the implicit dependence on $\{\mathbf{R_A}\}$ for simplicity.

A straightforward idea is to expand the many-electron wavefunction $\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$ using some mathematical expression of a set of one-electron spin orbitals $\{\chi(\mathbf{x})\}$, where $\chi(\mathbf{x}) = \psi(\mathbf{r})\alpha(\omega)$ for spin-up electrons and $\chi(\mathbf{x}) = \psi(\mathbf{r})\beta(\omega)$ for spin-down electrons. Here, $\{\psi(\mathbf{r})\}$ is the one-electron space orbitals. The expansion of the wavefunction $\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$ should at least satisfy two requirement:

1. All electrons are mathematically indistinguishable.

2. The many-electron wavefunction must be antisymmetric with repect to the interchange of the coordinate $\mathbf{x}$ of any two electrons [31]:

$$\Psi(\mathbf{x_1}, ..., \mathbf{x_i}, ..., \mathbf{x_j}, ..., \mathbf{x_N}) = -\Psi(\mathbf{x_1}, ..., \mathbf{x_j}, ..., \mathbf{x_i}, ..., \mathbf{x_N}). \tag{1.9}$$

For this purpose, the antisymmetric N-electron wavefunction can be expanded using the determinant:

$$\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x_1}) & \chi_2(\mathbf{x_1}) & \cdots & \chi_N(\mathbf{x_1}) \\ \chi_1(\mathbf{x_2}) & \chi_2(\mathbf{x_2}) & \cdots & \chi_N(\mathbf{x_2}) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x_N}) & \chi_2(\mathbf{x_N}) & \cdots & \chi_N(\mathbf{x_N}) \end{vmatrix}, \tag{1.10}$$

which is called a Slater determinant [32]. The Slater determinant has N electrons occupying N spin orbitals without specifying which electron is in which orbital, and is often shortened

as $|\chi_1\chi_2\cdots\chi_N\rangle$. Note that normally orthonormal spin orbitals $\{\chi(\mathbf{x})\}$ are used to construct the Slater determinant. It provides the basis for representing the many-electron wavefunction in most of the popular quantum chemistry methods.

The Hartree-Fock method assumes that the exact N-electron wavefunction of the system can be approximated by a single Slater determinant. One can then apply the variational theorem to minimize the ground-state energy $\langle\chi_1\chi_2\cdots\chi_N|\hat{H}|\chi_1\chi_2\cdots\chi_N\rangle$ with respect to the spin orbitals $\{\chi(\mathbf{x})\}$ to determine the best set of spin orbitals within the single determinant approximation. With the orthonormal constraint for spin orbitals $\{\chi(\mathbf{x})\}$, one can construct a Lagrangian

$$L[\{\chi(\mathbf{x})\}] = \langle\chi_1\chi_2\cdots\chi_N|\hat{H}|\chi_1\chi_2\cdots\chi_N\rangle - \sum_{ij}\epsilon_{ij}(\langle\chi_i|\chi_j\rangle - \delta_{ij}), \tag{1.11}$$

where ϵ_{ij} are the undetermined Lagrange multipliers. Take the first variation $\delta L = 0$, one arrives at the Hartree-Fock equations:

$$h(\mathbf{x_1})\chi_i(\mathbf{x_1}) + \sum_{j\neq i}\Big[\int d\mathbf{x_2}|\chi_j(\mathbf{x_2})|^2 r_{12}^{-1}\Big]\chi_i(\mathbf{x_1}) - \sum_{j\neq i}\Big[\int d\mathbf{x_2}\chi_j^*(\mathbf{x_2})\chi_i(\mathbf{x_2})r_{12}^{-1}\Big]\chi_j(\mathbf{x_1}) = \epsilon_i\chi_i(\mathbf{x_1}). \tag{1.12}$$

$h(\mathbf{x_1})$ is the one-particle Hamiltonian which includes the potential energy from the attraction of nuclei and kinetic energy of a single electron. The second term in Eq. 1.12 describes the Coulomb interaction of an electron in spin orbital χ_i with the average charge distribution of other $N-1$ electrons in the other spin orbitals. Here, we can define a Coulomb operator as:

$$J_j(\mathbf{x_1}) = \int d\mathbf{x_2}|\chi_j(\mathbf{x_2})|^2 r_{12}^{-1}, \tag{1.13}$$

which represents the average local potential at $\mathbf{x_1}$ arising from an electron in χ_j. The third term in Eq. 1.12 is called the exchange term and arises from the antisymmetric nature of the singlet determinant. We can also define an exchange operator $K_j(\mathbf{x_1})$ by expressing its operation on a spin orbital χ_i:

$$K_j(\mathbf{x_1})\chi_i(\mathbf{x_1}) = \sum_{j\neq i}\Big[\int d\mathbf{x_2}\chi_j^*(\mathbf{x_2})\chi_i(\mathbf{x_2})r_{12}^{-1}\Big]\chi_j(\mathbf{x_1}). \tag{1.14}$$

Inserting Eq. 1.13 and Eq. 1.14 into Eq. 1.12, and realizing

$$\Big[J_i(\mathbf{x_1}) - K_i(\mathbf{x_1})\Big]\chi_i(\mathbf{x_1}) = 0, \tag{1.15}$$

we obtain a more concise Hartree-Fock equation:

$$\Big[h(\mathbf{x_1}) + \sum_j J_j(\mathbf{x_1}) - \sum_j K_j(\mathbf{x_1})\Big]\chi_i(\mathbf{x_1}) = \epsilon_i\chi_i(\mathbf{x_1}). \tag{1.16}$$

We can further define a new operator as:

$$f(\mathbf{x_1}) = h(\mathbf{x_1}) + \sum_j J_j(\mathbf{x_1}) - \sum_j K_j(\mathbf{x_1}), \tag{1.17}$$

where $f(\mathbf{x_1})$ is called the Fock operator. The Hartree-Fock expression is further simplified as:

$$f(\mathbf{x_1})\chi_i(\mathbf{x_1}) = \epsilon_i\chi_i(\mathbf{x_1}). \tag{1.18}$$

Now it becomes clear that the Hartree-Fock method is a mean-field method, which means instead of explicit Coulomb interactions, each electron only feels an average electric field generated by its surrounding electrons.

To solve for the Hartree-Fock equation in Eq. 1.18, we introduce a set of atomic orbital basis functions $\{\phi_\mu\}$. Then we can expand the spin orbitals using

$$\chi_i = \sum_\mu C_{\mu i}\phi_\mu. \tag{1.19}$$

Insert Eq. 1.19 into Eq. 1.18, left multiply by ϕ_μ^* and integrate the equation, we arrive at

$$\sum_\nu C_{\nu i} \langle\phi_\mu| f |\phi_\nu\rangle = \epsilon_i \sum_\nu C_{\nu i} \langle\phi_\mu|\phi_\nu\rangle, \tag{1.20}$$

which is also called the Roothaan equation. Introducing the matrix element notation

$$\begin{aligned} S_{\mu\nu} &= \langle\phi_\mu|\phi_\nu\rangle \\ F_{\mu\nu} &= \langle\phi_\mu| f |\phi_\nu\rangle, \end{aligned} \tag{1.21}$$

we obtain the final working equation for the Hartree-Fock method:

$$\mathbf{FC} = \mathbf{SC}\epsilon, \tag{1.22}$$

where $\mathbf{F}$ is the Fock matrix, $\mathbf{S}$ is the overlap matrix, ϵ and $\mathbf{C}$ are the eigenvalue and eigenvector matrices. By using some linear algebra tricks, it turns out that solving Eq. 1.22 is equal to diagonalizing the Fock matrix $\mathbf{F}$. However, one may notice that the Fock matrix $\mathbf{F}$ depends on the eigenvector matrix $\mathbf{C}$, which is the solution of diagonalizing $\mathbf{F}$. Therefore, one needs to start with an initial guess of $\mathbf{C}$ and perform Eq. 1.22 iteratively until the solution $\mathbf{C}$ converges. Because of this process, Hartree-Fock is often referred to as the self-consistent field (SCF) method.

Hartree-Fock is a crude approximation that causes large errors in both electronic wavefunction and energy. The mean-field approximation, which replaces the instantaneous electron-electron repulsion with the repulsion of each electron with an average electron charge cloud, makes the electron correlation to be completely missing in the Hartree-Fock method. Meanwhile, Hartree-Fock assumes the wavefunction can be written as a single Slater determinant, which is also an inaccurate approximation in many cases. Thus, many wavefunction methods have been developed based on Hartree-Fock, which are called post-Hartree-Fock methods, in order to capture the missing electron correlation in Hartree-Fock.

1.1.2 Configuration interaction

In Hartree-Fock theory, one assumes the antisymmetric wavefunction of the system can be expanded using a single Slater determinant, which is not true. In theory, an arbitrary N-electron wavefunction can only be expressed exactly as a linear combination of all possible N-electron basis functions formed from a complete set of one-electron basis functions. In reality, the N-electron basis functions can be chosen as N-electron Slater determinants and one-electron basis functions are spin orbitals $\{\chi_i(\mathbf{x})\}$. Hartree-Fock just picks one of all possible N-electron Slater determinant for the N-electron wavefunction expansion – no wonder it performs poorly in many chemical systems.

The idea of using all possible Slater determinants for the wavefunction expansion leads to the development of the configuration interaction (CI) method [33]. For simplicity, suppose we have already solved the Hartree-Fock equation for an N-electron (N is an even number)

system in a finite one-electron basis set $\{\phi_\mu\}$ and obtained a set of $2K$ spin orbitals $\{\chi_i\}$. The determinant formed from the N lowest energy spin orbitals is the HF determinant Φ_0. Meanwhile, we can also form a large number of other N-electron determinants from the $2K$ spin orbitals. For example, instead of using the occupied HF orbital χ_a, we can use an unoccupied HF orbital χ_r in the determinant, leading to a new determinant Φ_a^r. The notation Φ_a^r means we "excite" one electron from occupied HF orbital χ_a to unoccupied HF orbital χ_r, and this is called a singly excited determinant. Similarly, a doubly excited determinant χ_{ab}^{rs} can be formed by replacing the occupied HF orbitals χ_a and χ_b with unoccupied HF orbitals χ_r and χ_s. In this way, we can form up to N-tuply excited determinants so that we have formed all possible N-electron Slater determinants in the given finite one-electron basis set. These Slater determinants are often referred to as "configurations".

Using these N-electron Slater determinants, we can expand the electronic wavefunction Ψ as:

$$\Psi = c_0\Phi_0 + \sum_{ar} c_a^r \Phi_a^r + \sum_{\substack{a<b \\ r<s}} c_{ab}^{rs} \Phi_{ab}^{rs} + \sum_{\substack{a<b<c \\ r<s<t}} c_{abc}^{rst} \Phi_{abc}^{rst} + \cdots , \tag{1.23}$$

where c_0, c_a^r and c_{ab}^{rs} are the expansion coefficients for the HF, singly excited and doubly excited determinants. Eq. 1.23 is a called a full configuration interaction (FCI) expansion, which includes all possible configurations to expand the wavefunction Ψ. Using this expansion, one can then apply the variational theorem to minimize the ground-state energy $\langle\Psi|\hat{H}|\Psi\rangle$ with respect to the CI coefficients $\{c_0, \{c_a^r\}, \{c_{ab}^{rs}\}, ...\}$. This method is called full configuration interaction method, which can be seen as a direct matrix mechanics solution to the electronic Schrödinger equation. Note that we use a finite one-particle basis set $\{\phi_\mu\}$ here, so Eq. 1.23 is still not an exact expansion of the wavefunction Ψ, which means the FCI solution is not the exact solution to the Schrödinger equation. Ideally, if we adopt an infinitely large one-particle basis set and employ all possible determinants, we can reach an exact solution, which is called a complete CI. In practice, we always need to use a finite one-particle basis and FCI is the best solution we can obtain within the given basis set.

FCI can be applied to accurately compute not only the ground state, but also the excited states of both closed-shell and open-shell molecules. The difference between the FCI ground-state energy and the HF energy defines the electron correlation energy:

$$E_{corr} = E^{FCI} - E^{HF}. \tag{1.24}$$

Unfortunately, the FCI method is too expensive to be solved because there are too many Slater determinants involved. Employing configuration state functions or efficient algorithms can speed up the FCI calculations, but generally FCI still has a factorial scaling with respect to the system size. This high computational scaling limits FCI to be only applied to small-sized molecules, which makes it to serve as a benchmark method to evaluate the performance of other quantum chemistry methods.

Truncating the FCI expansion is another idea to enable efficient CI calculations. Realizing that the Hamiltonian is a combination of one-particle and two-particle operators, one assumes that the singly and doubly excited determinants are the most important configurations, because they are the only determinants that can directly interact with the reference HF determinant Φ_0. This approximation leads to methods such as the CI singles (CIS) and CI singles and doubles (CISD). It is indeed a good approximation as it reduces the number of configurations significantly while can still account for about 95% of the electron correlation in CISD [33]. However, the CI truncation also brings errors. One most severe problem is that truncated CI methods are not size extensive, which means the energy calculated for non-interacting particles does not scale linearly with the particle numbers.

In this thesis, we employ the FCI method as the high-level theory to correct the errors in density functional theory. Density decomposition methods are developed so that the FCI method is still affordable for the studied systems. In addition to CI, many other quantum chemistry methods have also been developed, among which the most commonly-used ones are the many-body perturbation theory [34] and coupled-cluster theory [35]. Especially, coupled-cluster singles and doubles with perturbative triples (CCSD(T)) [36] is now the "gold standard" for computing weakly-correlated chemical systems. CCSD(T) results are used as a reference to test the performance of our methods on thermochemistry in this thesis. We will not further review these methods but now turn our attention to the density functional theory.

1.2 Density functional theory

In Section 1.1, all quantum chemistry methods use the wavefunction $\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$ as the key quantity to describe the system. In the density functional theory (DFT), the key quantity becomes the electron density $\rho(\mathbf{r})$, which is defined as the integral over the spin

coordinates of all electrons and over all but one of the spatial variables:

$$\rho(\mathbf{r}) = N \int \cdots \int \left|\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})\right|^2 d\omega_1 d\mathbf{x_2} ... d\mathbf{x_N}. \tag{1.25}$$

The physical meaning of $\rho(\mathbf{r})$ is that it determines the probability of finding any of the N electrons within volume $d\mathbf{r}$. The wavefunction $\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N})$ of an N-electron system includes $3N$ space variables, while the electron density $\rho(\mathbf{r})$ always only has three space variables. Therefore, employing the electron density as the basic quantity reduces the complexity in the wavefunction significantly and may lead to much lower computational scaling. In addition, unlike the wavefunction, the electron density is a variable that can be measured experimentally, so it can be used to directly compare the theory with experiments.

In 1964, Hohenberg and Kohn established the foundation for the density functional theory [37]. They stated in the first Hohenberg-Kohn theorem that, the external potential, and hence the total energy, is a unique functional of the electron density. This theorem proves that the electron density $\rho(\mathbf{r})$ uniquely determines the Hamiltonian operator (i.e., the external potential $v_{ext}(\mathbf{r})$) and thus all the properties of the system. Thus, the electron density $\rho(\mathbf{r})$ can be used as the fundamental variable in quantum chemistry. The second Hohenberg-Kohn theorem states that, the functional that delivers the ground state energy of the system, gives the lowest energy if and only if the input density is the true ground state density. This theorem provides the variational principle for the density functional theory. It also indicates that the electron density in DFT needs to be "v-representable", which means $\rho(\mathbf{r})$ has to be the corresponding density of some external potential $v_{ext}(\mathbf{r})$.

Hohenberg-Kohn theorems guarantee that there exists a universal functional $F[\rho(\mathbf{r})]$, that predicts the electronic ground state of all molecules and materials exactly. However, knowing the existence of such a functional does not mean we can also write it down. Actually, it is the birth of the Kohn-Sham theory that puts DFT into the successful practical use.

1.2.1 Kohn-Sham density functional theory

One of the most difficult tasks in DFT is to design an accurate functional to compute the kinetic energy. The Thomas-Fermi model [38, 39] is one such approximate functional:

$$T^{TF}[\rho(\mathbf{r})] = \frac{3}{10}(3\pi^2)^{2/3} \int \rho(\mathbf{r})^{5/3} d\mathbf{r}, \tag{1.26}$$

but all such approximations produce large errors for kinetic energies of atoms and molecules.

In 1965, Kohn and Sham designed a nice framework to deal with this problem [21]. The Kohn-Sham density functional theory (KS-DFT) assume there is a fictitious system of non-interacting electrons that has the same density as our interested system of interacting electrons. This non-interacting fictitious system can be represented using a single Slater determinant of orthonormal orbitals $\{\psi_i\}$, just as in the Hartree-Fock theory. So the electron density can be represented as:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2. \tag{1.27}$$

Then the kinetic energy of the interacting electrons can be approximately computed as the kinetic energy of the non-interacting fictitious system:

$$T[\rho(\mathbf{r})] \approx T_s[\{\psi_i\}] = -\frac{1}{2}\sum_{i=1}^{N} \int \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r}. \tag{1.28}$$

The nucleus-electron attraction energy is

$$V_{ne}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})v_N(\mathbf{r})d\mathbf{r}, \tag{1.29}$$

where $v_N(\mathbf{r})$ is the nuclear potential, and the electron-electron repulsion energy $V_{ee}[\rho(\mathbf{r})]$ can be approximated by the classical Coulomb self energy

$$J[\rho(\mathbf{r})] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r_1})\rho(\mathbf{r_2})}{r_{12}} d\mathbf{r_1}d\mathbf{r_2}. \tag{1.30}$$

Thus, the KS-DFT total energy of the interacting system is expressed as

$$E[\rho(\mathbf{r})] = T_s[\{\psi_i\}] + V_{ne}[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})], \tag{1.31}$$

where the last term $E_{xc}[\rho(\mathbf{r})]$ is called the exchange-correlation (XC) energy. The XC energy functional describes all the errors made by the approximations above:

$$E_{xc}[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] - T_s[\{\psi_i\}] + V_{ee}[\rho(\mathbf{r})] - J[\rho(\mathbf{r})]. \tag{1.32}$$

If we know the exact form of the XC functional, we would solve the Schrödinger equation exactly. Thus, KS-DFT is also a formally exact theory. However, in reality, the analytic

form of the XC functional is never known and one always has to approximate this term. Compared to the orbital-free DFT, the KS-DFT energy is an implicit density functional that explicitly depends on the KS orbitals. However, the strength of such scheme is that, it captures most of the contributions of the kinetic energy and electron-electron repulsion through $T_s[\{\psi_i\}]$ and $J[\rho(\mathbf{r})]$, and one just needs to find a good approximation to the last piece – XC functional.

Take the variational minimization of Eq. 1.31 with respect to the orbital $\psi_i(\mathbf{r})$ yields the Kohn-Sham equation:

$$\Big(-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\Big)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \tag{1.33}$$

where $v_s(\mathbf{r})$ is the effective potential

$$v_s(\mathbf{r}) = v_N(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}), \tag{1.34}$$

and $v_{xc}(\mathbf{r})$ is the exchange-correlation potential

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho}. \tag{1.35}$$

Solving Eq. 1.33 employs a similar self-consistent field algorithm as in the Hartree-Fock method. KS-DFT is operationally an independent-particle theory and is as simple as, if not simpler than, the Hartree-Fock theory. But it delivers the exact electron density and other electronic properties in principle. KS-DFT with a good approximate XC functional can describe the electron correlation much better than HF. Thus, most of the development in the field of DFT has been focused on searching for better XC functionals since the foundation of KS-DFT.

1.2.2 Jacob's ladder of density functional approximations

In this section, we review some of the efforts made for developing XC functionals. The central idea of designing approximate XC functionals is that one hopes to achieve better accuracy by using more and more complicated design ingredients in the functionals. This idea is summarized by John Perdew's Jacob's ladder of density functional approximations [40, 41], as shown in Figure 1-1. According to the design ingredients used in the functional, one puts the XC functionals onto different rungs of this ladder, and expects the functionals on the

higher rungs have better performance than the lower ones.

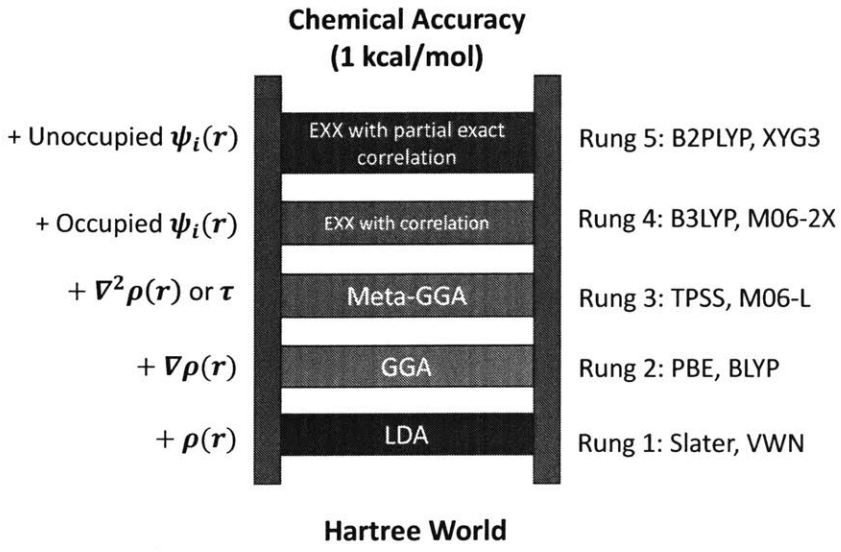


Figure 1-1: John Perdew's Jacob's ladder of density functional approximations to the exchange-correlation energy.

We start from the "Hartree world" [42], where no XC functional is used in the KS-DFT (Eq. 1.31). Our final goal of density functional development is to be able to describe the thermochemistry within the accuracy of 1 kcal/mol, which is called the "chemical accuracy". The simplest density functionals depend only on the electron density $\rho(\mathbf{r})$, which are called the local density approximation (LDA), or local spin-density approximation (LSDA) [40]:

$$E_{xc}^{LSDA}[\rho_\alpha, \rho_\beta] = \int \rho(\mathbf{r})\epsilon_{xc}(\rho_\alpha, \rho_\beta)d\mathbf{r}, \tag{1.36}$$

where the ρ_α and ρ_β are the spin-up and spin-down electron densities and $\epsilon_{xc}(\rho_\alpha, \rho_\beta)$ is the exchange-correlation energy per electron of an uniform electron gas of density $(\rho_\alpha, \rho_\beta)$. LDA (LSDA) functionals define Rung 1 of the Jacob's ladder. An example of the LSDA approximation is the Slater exchange functional [22]:

$$E_x^{Slater}[\rho_\alpha, \rho_\beta] = -\frac{3}{2}(\frac{3}{4\pi})^{1/3}\sum_{\sigma}^{\alpha,\beta}\int \rho_\sigma^{4/3}d\mathbf{r}. \tag{1.37}$$

LDA (LSDA) generally tends to underestimate the exchange energy and overestimate the correlation energy. It assumes the density is the same everywhere, which is a poor estimation

especially for molecules where the electron density changes rapidly in many regions.

It is therefore natural to introduce an ingredient that can account for inhomogeneities in the density, which is the electron density gradient $\nabla\rho(\mathbf{r})$. This gives rise to the generalized gradient approximation (GGA), which is located on Rung 2 of the Jacob's ladder [40]:

$$E_{xc}^{GGA}[\rho_\alpha, \rho_\beta] = \int \rho(\mathbf{r})\epsilon_{xc}(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta)d\mathbf{r}. \tag{1.38}$$

The GGA functionals show great improvement over the LDA functionals, predicting good results for molecular geometries and ground-state energies. Several popular GGA functionals are PBE [43], BLYP [44, 45] and PW91 [46], among which PBE is still the most commonly-used functional in materials science nowadays.

The improvement from LDA to GGA suggests the idea of adding more design ingredients works well. Following this idea, one can add one more ingredient, the Laplacian of electron density $\nabla^2\rho(\mathbf{r})$ or the kinetic energy density $\tau = \frac{1}{2}\sum_i |\nabla\psi_i|^2$, which capture the second derivative information of the density. This leads to the third rung, meta-GGA functionals [40]:

$$E_{xc}^{mGGA}[\rho_\alpha, \rho_\beta] = \int \rho(\mathbf{r})\epsilon_{xc}(\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta, \nabla^2\rho_\alpha, , \nabla^2\rho_\beta, \tau_\alpha, \tau_\beta)d\mathbf{r}. \tag{1.39}$$

Adding the second derivative information indeed sometimes results in better performance. For example, compared to PBE (GGA), TPSS (meta-GGA) [47] greatly improves atomization energies for molecules and surface energies for solids [23].

Another route to design more accurate density functionals is to add ingredients of KS orbitals. First such development is the introduction of HF exchange into the XC functionals, which is also called the exact exchange (EXX) functional in DFT:

$$E_x^{EXX} = -\frac{1}{2}\sum_{ij}^{occ}\int\int \frac{\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'. \tag{1.40}$$

The mixture of HF exchange and GGA (and meta-GGA) functionals results in the hybrid functionals (Rung 4 on the Jacob's ladder), where the percentage of HF exchange can be determined non-empirically [48] or by fitting to reproduce accurate thermochemistry data [49].

The latter way leads to the construction of the famous B3LYP functional [49, 50]:

$$E_{xc}^{B3LYP} = c_x E_x^{HF} + (1 - c_x - a_x) E_x^{Slater} + a_x E_x^{B88} + (1 - a_c) E_c^{VWN} + a_c E_c^{LYP}, \quad (1.41)$$

where $c_x = 0.20$, $a_x = 0.72$ and $a_c = 0.81$. The inclusion of the HF exchange help significantly reduce some errors (one of them is called self-interaction error) in GGA (meta-GGA) functionals, which we will further discuss in later chapters. Such correction leads to very successful applications of KS-DFT to chemistry problems and the hybrid functionals are especially good at predicting molecular structures and reaction energies. To date, B3LYP is still the most widely used functional in DFT applications.

The fifth rung on the Jacob's ladder utilizes unoccupied KS orbitals. For example, the second-order perturbation theory (PT2) correlation energy using KS orbitals and eigenvalues [51] may be partially used in the construction of correlation functionals, such as in B2PLYP [52] and XYG3 [53] functionals. Another type of functionals [54, 55] use the idea of the fluctuation dissipation theorem [56, 57] and thus employ the random phase approximation (RPA) correlation energy.

Overall speaking, the idea of constructing density functionals according to the Jacob's ladder shows promise towards better density functionals. However, problems exist for such a density functional hierarchy. The most severe problem is, in many cases, the functionals on higher rungs are not necessarily better than the ones on lower rungs [23]. For particular applications, it is entirely possible that PBE (Rung 2) is more accurate than PBE0 (Rung 4). Such uncertainty is what makes DFT less reliable than wavefunction theory, in which one almost always obtain better accuracy by inputting more computational power. This is partly because common density functionals have serious systematic errors, such as the self-interaction error, inability to describe strong correlations or dispersion interactions [23]. In reality, one almost always needs to carefully benchmark the performance of approximate functionals before drawing any reliable conclusions from KS-DFT calculations. Thus, the biggest challenge in density functional development is probably to design a hierarchy that can systematically improve the performance of approximate DFT calculations, which is also the main goal of this thesis.

1.3 Unconventional DFT methods

In this section, we review some of the unconventional DFT methods that either provide the basis for our new DFT method development or serve as a tool for applications in OLED simulations in this thesis. These methods are unconventional in the sense that they are not regular ground-state or excited-state KS-DFT methods, but are especially efficient at dealing with some particular chemical problems.

1.3.1 Restricted open-shell Kohn-Sham

In this section, we review an unconventional KS-DFT based excited-state method – restricted open-shell Kohn-Sham (ROKS) [58, 59]. Accurate methods for modeling electronically excited states are playing an increasingly important role in the design of optoelectronic materials including photovoltaics [14], light-emitting diodes [16] and field-effect transistors [60]. Many of the molecules of interest are too large (>100 atoms) to be studied with high-level wavefunction-based methods like CASPT2 [61] or coupled-cluster [7], which are known to give quite accurate results for small organic molecules [62]. As a result, DFT-based approaches are the only viable post-Hartree-Fock computational tools for studying such molecules. The most standard DFT-based excited state method is linear response time-dependent density functional theory (TDDFT) [63, 64]. Despite the dominance of TDDFT in organic materials simulation, ROKS is a similarly reliable excited-state method and performs better than TDDFT in some applications involving charge-transfer excited states.

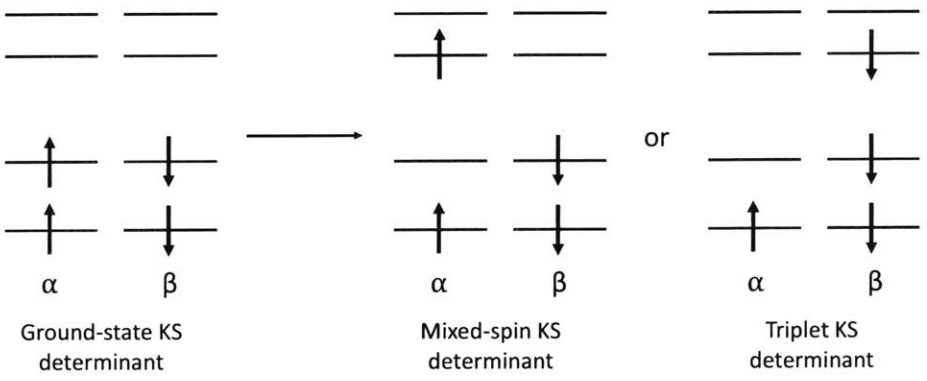


Figure 1-2: Kohn-Sham excited-state determinants used to represent ROKS S_1 state.

ROKS operates as a ground-state KS-DFT approach to access the lowest singlet excited

state (S_1). In many optoelectronic applications, the lowest singlet excited state energy is primarily concerned. The central idea of ROKS is to follow a non-Aufbau rule of occupying electrons when solving a KS-DFT equation. Instead of occupying all electrons on lowest-energy spin orbitals, one electron is "excited" to the lowest unoccupied orbital, leading to two different kinds of restricted open-shell determinants: a mixed-spin KS determinant and a triplet KS determinant, as shown in Figure 1-2. The S_1 state can be represented using these two determinants, whose energy is expressed as [58]:

$$E_s^{ROKS} = 2E_m[\{\phi_i\}] - E_t[\{\phi_i\}], \tag{1.42}$$

where s, m and t denote the singlet excited state, mixed-spin determinant, and triplet determinant, respectively. Variational minimization of the ROKS energy in Eq. 1.42 with respect to the KS orbitals leads to an eigenvalue equation:

$$\mathbf{FC} = \mathbf{C}\epsilon, \tag{1.43}$$

but the Fock matrix $\mathbf{F}$ is much more complicated than a normal KS-DFT Fock matrix, which we will not further discuss.

ROKS has been shown to be competitive with TDDFT for excited-state energies and geometries of organic molecules when hybrid and range-separated functionals are used [59]. Particularly, since ROKS operates like a ground-state method, it may not share the problem of TDDFT to significantly underestimate the charge-transfer excited states. Therefore, in this thesis, we employ ROKS to study the excited states of OLED molecules. A special type of the studied OLED molecules is the thermally activated delayed fluorescence (TADF) OLED, which possesses charge-transfer excited states and small singlet-triplet energy gap.

1.3.2 Constrained density functional theory

In this section, we review an unconventional KS-DFT method to treat charge transfer (CT) problems – constrained density functional theory (CDFT) [65, 66, 67]. Marcus theory [68, 69] provides a closed form expression for the charge transfer rate. Its practical use hinges on the possibility of calculating diabatic charge-localized states [70, 71] which have similar electronic configurations at all nuclear coordinates. For instance, in a charge separation process, an electron transfers from the donor (D) to the acceptor (A); hence, DA is one

diabatic state and D^+A^- is the other. Standard KS-DFT methods diagonalize the effective electronic Hamiltonian; therefore, they generate adiabatic states that are approximations to the true eigenstates of the Hamiltonian for the D-A system.

To obtain an charge-localized diabatic state, one wants to explicitly constrain the electron density to the corresponding configuration (e.g., D^+A^- or DA) and perform all calculations under the constraint. This is achieved in constrained density functional theory (CDFT) through minimization of the total energy functional under an additional constraint that the charge densities of fragments integrate to the predefined numbers of electrons. Since there is no unique way to partition the electron density into fragments, a weight function $w_c(\mathbf{r})$ is introduced to define a charge constraint on the chosen fragment (may be D or A):

$$\int w_c(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} = N_c. \tag{1.44}$$

While there is much freedom in the definition of the weight function $w_c(\mathbf{r})$, in the majority of implementations it is based on Becke grid weights [72] or Hirshfeld density partitioning [73]. Despite this ambiguity, the electronic couplings computed with CDFT states show only limited sensitivity to the choice of the weight function [74]. Then, a constraining potential V_c is introduced as the Lagrange multiplier so that minimizing the energy of the systems subject to the constraint in Eq. 1.44 is equivalent to finding the stationary point of

$$W[\rho, V_c] = E[\rho] + V_c\Big(\int w_c(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + N_c\Big) \tag{1.45}$$

with respect to ρ and V_c. This leads to Kohn-Sham equations that need to be solved to obtain a KS charge-localized state:

$$\Big(-\frac{1}{2}\nabla^2 + v_N(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}) + V_c w_c(\mathbf{r})\Big)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \tag{1.46}$$

where $V_c w_c(\mathbf{r})$ is the constraining potential that enforces the proper charge on the chosen fragment. Since the optimal Lagrange multiplier V_c is not known a priori, it has to be optimized together with the electron density ρ. In practice, this is done by adding an extra optimization loop to enforce Eq. 1.44 at each SCF cycle [75].

The CDFT diabatic states can then be used to compute useful quantities in Marcus theory [66]. For example, the free energy difference between two diabatic states DA and

D^+A^- and the inner-sphere reorganization energy can be computed as:

$$\Delta G = E(D^+A^-|D^+A^-) - E(DA|DA), \tag{1.47}$$

$$\lambda_i = E(D^+A^-|DA) - E(D^+A^-|D^+A^-), \tag{1.48}$$

where $E(a|b)$ means the energy of state a calculated at the equilibrium structure of state b. In this thesis, we employ CDFT to investigate the charge transfer mechanisms in a condensed phase OLED simulation.

1.3.3 Potential inversion in KS-DFT

In this section, we review a technique that allows one to calculate KS kinetic energy and orbital-dependent exchange-correlation functional directly from the given electron density, which is called the potential inversion. In KS-DFT, the kinetic energy functional $T_s[\rho]$ is known not as an explicit functional of the electron density $\rho(\mathbf{r})$, but as an implicit functional of $\rho(\mathbf{r})$, through the dependence on one-electron KS orbitals $\{\phi_i(\mathbf{r})\}$. This dependence makes the kinetic energy an explicit functional of KS orbitals $T_s[\{\phi_i(\mathbf{r})\}]$. One has to first know the KS orbitals $\{\phi_i(\mathbf{r})\}$ to be able to compute $T_s[\{\phi_i(\mathbf{r})\}]$. Many exchange-correlation functionals are also implicit density functionals. One example is the exact exchange (EXX) functional defined in Eq. 1.40, which is again an explicit functional of KS orbitals $\{\phi_i(\mathbf{r})\}$ and very important for constructing hybrid functionals.

An interesting problem is that, if an electron density $\rho_{in}(\mathbf{r})$ is given a priori, how can one compute the implicit density functionals such as the KS kinetic energy $T_s[\rho_{in}]$? Such technique is called the potential inversion [76, 77, 78]. Instead of optimizing the electron density $\rho(\mathbf{r})$ in a standard KS-DFT equation, we try to search for the effective KS potential $v_s(\mathbf{r})$ (Eq. 1.34) that produces $\rho_{in}(\mathbf{r})$ through an eigenvalue equation. With the effective KS potential $v_s(\mathbf{r})$, one can then obtain the corresponding KS orbitals and compute all implicit density functionals.

The potential inversion technique is needed for two reasons. First, it would help the development of better explicit density functionals. Assume we can get the highly accurate density $\rho_{in}(\mathbf{r})$ from high-level wavefunction theory calculations (e.g. FCI, CCSD(T)) or experiments. Using the potential inversion method, one can obtain the highly accurate effective KS potential $v_s(\mathbf{r})$, and thus the exchange-correlation potential $v_{xc}(\mathbf{r})$ (Eq. 1.35).

Note that $v_{xc}(\mathbf{r})$ is coming from the functional derivative of $E_{xc}[\rho]$ with respect to the density $\rho(\mathbf{r})$. This highly accurate $v_{xc}(\mathbf{r})$ can therefore serve as a guidance to design more accurate XC functional $E_{xc}[\rho]$. Because of this reason, the potential inversion is often referred to as optimized effective potential (OEP) [79, 80, 81, 82]. Second, the potential inversion provides a framework for developing new fragment-based DFT methods. Fragment-based DFT methods have been developed recently due to its potential to provide faster and more accurate DFT simulations. If we manually partition the system into several fragments and want to treat these fragments separately, the potential inversion provides a way to directly compute the KS-DFT energy for a given fragment density. This technique has been extensively used in methods such as DFT embedding [83, 84] and partition DFT [85], and also lays the foundation for our many-pair expansion method in this thesis.

Now we review a numerical method for performing potential inversions in the KS-DFT framework and closely follow the algorithm proposed by Wu and Yang [77, 86]. For an N-electron spin-compensated system and a given input electron density $\rho_{in}(\mathbf{r})$, the KS kinetic energy is defined as [87, 22]:

$$T_s[\rho_{in}] = \min_{\{\phi_i\}\to\rho_{in}} 2\sum_{i}^{N/2} \langle\phi_i|\,\hat{T}\,|\phi_i\rangle\,, \tag{1.49}$$

where $\{\phi_i(\mathbf{r})\}$ are doubly-occupied KS orbitals that need to be searched through the kinetic energy minimization. The electron density $\rho(\mathbf{r})$ computed from the KS orbitals

$$\rho(\mathbf{r}) = 2\sum_{i}^{N/2} \left|\phi_i(\mathbf{r})\right|^2 \tag{1.50}$$

should be equal to the input electron density

$$\rho(\mathbf{r}) = \rho_{in}(\mathbf{r}). \tag{1.51}$$

Therefore, the evaluation of $T_s[\rho]$ is a constrained minimization problem. We introduce a potential $v(\mathbf{r})$ as the Lagrange multiplier and define the following functional of KS orbitals $\{\phi_i(\mathbf{r})\}$ and the potential $v(\mathbf{r})$:

$$W_s[\{\phi_i\}, v(\mathbf{r})] = 2\sum_{i}^{N/2} \langle\phi_i|\,\hat{T}\,|\phi_i\rangle + \int v(\mathbf{r})\Big(\rho(\mathbf{r}) - \rho_{in}(\mathbf{r})\Big)d\mathbf{r}. \tag{1.52}$$

Optimizing $W_s[\{\phi_i\}, v(\mathbf{r})]$ with respect to KS orbitals $\{\phi_i(\mathbf{r})\}$, one arrives at

$$\big(\hat{T} + v(\mathbf{r})\big)\phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \tag{1.53}$$

so $\{\phi_i(\mathbf{r})\}$ can be calculated as the eigenfunctions of the operator $\hat{T} + v(\mathbf{r})$.

In order to use standard linear algebra techniques, we need a way to represent the potential $v(\mathbf{r})$. Here, we expand $v(\mathbf{r})$ as:

$$v(\mathbf{r}) = v_N(\mathbf{r}) + v_0(\mathbf{r}) + \sum_t b_t g_t(\mathbf{r}), \tag{1.54}$$

where $v_N(\mathbf{r})$ is the nuclear potential of the system, $v_0(\mathbf{r})$ is a fixed reference potential, and $\{b_t\}$ are the coefficients for the linear combination of a set of finite basis functions $\{g_t(\mathbf{r})\}$. For atoms and molecules, $\{g_t(\mathbf{r})\}$ can be chosen as the standard Gaussian basis set or Coulomb potential of the Gaussian basis functions. For $v_0(\mathbf{r})$, the Fermi-Amaldi potential is suggested [76]:

$$v_0(\mathbf{r}) = \frac{N-1}{N} \int \frac{\rho_{in}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{1.55}$$

which provides the correct long-range behavior of the exact $v(\mathbf{r})$. Since $v_N(\mathbf{r})$ and $v_0(\mathbf{r})$ are both fixed potentials, optimizing $W_s[\{\phi_i\}, v(\mathbf{r})]$ with respect to $v(\mathbf{r})$ is equivalent to optimizing $W_s[\{\phi_i\}, v(\mathbf{r})]$ with respect to the coefficients $\{b_t\}$. The first derivative of $W_s[\{\phi_i\}, v(\mathbf{r})]$ with respect to $\{b_t\}$ can be computed as:

$$\frac{\partial W_s[\{\phi_i\}, v(\mathbf{r})]}{\partial b_t} = \int \big(\rho(\mathbf{r}) - \rho_{in}(\mathbf{r})\big) g_t(\mathbf{r}) d\mathbf{r}, \tag{1.56}$$

and the second derivative is

$$\frac{\partial^2 W_s[\{\phi_i\}, v(\mathbf{r})]}{\partial b_t \partial b_u} = 2 \sum_i^{occ} \sum_a^{unocc} \frac{\langle \phi_i | \, g_u(\mathbf{r}) \, | \phi_a \rangle \, \langle \phi_a | \, g_t(\mathbf{r}) \, | \phi_i \rangle}{\epsilon_i - \epsilon_a} + c.c. \tag{1.57}$$

using Eq. 1.52 and Eq. 1.54. In this way, the constrained minimization of $T_s[\rho]$ is turned into the unconstrained maximization of $W_s[\{\phi_i\}, v(\mathbf{r})]$. With the first and second analytic derivatives, one can employ a standard Newton optimization method to find the converged coefficients $\{b_t\}$ and thus the effective potential $v(\mathbf{r})$. With the effective potential $v(\mathbf{r})$, KS orbitals and other implicit exchange-correlation functionals can also be obtained.

In this thesis, the potential inversion technique serves as a critical tool for our MPE

method development. We will further discuss the practical use of this tool in the later chapters.

1.4 Incorporating environment effects

One important goal of this thesis is to properly use available DFT methods for efficient modeling of electronic properties of complex chemical systems. One such chemical system is organic light-emitting diodes (OLEDs) [27] and we are particularly interested in the charge transfer (CT) properties of OLEDs. CT reactions usually occur in condensed phases, and the environment can have a significant effect on the electronic structure of reactants and the resulting reaction mechanisms. In fact, the environment's degrees of freedom play a crucial role in the CT process; their fluctuations make two diabatic states transiently isoenergetic, at which point the electron hop happens. On average the environment stabilizes localized charges through polarization of its electron density, alignment of multipole moments and geometric deformation (polaron formation). For these reasons accounting for the presence of the environment in DFT simulations of CT is of fundamental importance.

Unfortunately, although DFT is already an efficient electronic structure method, it is still impossible to carry out full DFT simulations of large-scale (>10000 atoms) condensed phase chemical systems. In practice, we only perform DFT calculations for the molecules involved in the charge transfer reactions, which are embedded in the environment described by more crude models. The methods for including embedding effects can be divided into explicit and implicit models, depending whether surrounding molecules are present explicitly in the simulations or are replaced by a dielectric continuous medium.

In Marcus theory [68, 69] for CT reactions, two thermodynamic properties need to be evaluated – the driving force and the reorganization energy. The driving force is the difference between the free energies of the initial and final diabatic states, so the role of the environment model is to describe the stabilization of the charges after the environment's degrees of freedom have been equilibrated. Calculations of the reorganization energies are more complicated as it is a non-equilibrium property, which requires calculations of the energy of one diabatic state in the equilibrium structure of the other. It is particularly challenging for the outer sphere component of the reorganization energy, which is associated with the rearrangement of surrounding molecules and its treatment depends on the type of

the model used. In this section, we briefly review the most commonly used implicit and explicit environment models for condensed phase simulations.

1.4.1 Implicit environment model

In implicit environment models [88, 89], the molecule or the reactive complex is placed inside a cavity resembling its molecular shape. The cavity is constructed either as a union of van der Waals spheres or is enclosed inside an isosurface of the molecule's electron density. All the space outside of the cavity is filled with a continuous medium characterized by its dielectric constant, which is usually assumed to be frequency independent. The presence of the molecular system inside the cavity induces charges in the dielectric, which generate an effective reaction field interacting with the embedded electron density. The free energy of interactions with the environment can be calculated by solving the Poisson equation for the apparent charge on the surface of the cavity. The surface charge $\sigma(\mathbf{r_s})$ generates a Coulomb potential

$$\phi_\sigma(\mathbf{r}) = \int \frac{\sigma(\mathbf{r_s})}{|\mathbf{r} - \mathbf{r_s}|} d\mathbf{r_s}, \tag{1.58}$$

which is then added to the Hamiltonian of the embedded molecule. In the self-consistent reaction field (SCRF) [90] family of approaches, the equations for the electronic structure of the embedded system and the surface charge on the cavity are solved iteratively until convergence. The converged electron density and electrostatic potential are used to compute the electrostatic contribution to the embedding free energy

$$\Delta G_{es} = \frac{1}{2} \int \phi_{reac}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}, \tag{1.59}$$

where the reaction field $\phi_{reac}(\mathbf{r})$ is the difference between electrostatic potentials of a molecule in the cavity and vacuum. Apart from the electrostatic component, the free energy of cavitation and dispersion/repulsion interactions can be accounted for in implicit models. These contributions are usually highly parametrized to fit to experimental or other theoretical data.

The implicit models used most frequently in computational studies of charge transfer reactions in condensed phases are the polarizable continuum model (PCM) [91] in its several variants [92], conductor-like screening model (COSMO) [93, 94] and the SMx family of models [95, 96], where the latter are based on solution of the generalized Born instead of the

Poisson equation. The common advantage of these methods is that they are computationally very efficient and account for the majority of important effects resulting from interaction with the environment. However, they rely on many fitted parameters, so transferability and accuracy are often an issue. Also, the specific interactions between the molecules, e.g., hydrogen bonding, are not accounted for in continuum models. Another challenge of modeling of CT reactions with implicit environment is that they assume thermodynamic equilibrium and there is no direct way to freeze environment's geometric relaxation in calculations of reorganization energies. Extensions of implicit models to non-equilibrium structures assume that the bulk dielectric constant has a slow and fast (ϵ_∞) component [97], where the latter is associated with electronic polarization only (high-frequency component). Then, ϵ_∞ can be used in calculations of reorganization energies, which is a qualitative yet efficient approach.

1.4.2 Explicit environment model

While implicit models are efficient, they necessarily ignore the fact that the environment is composed of individual molecules. Conversely, in explicit environment models the molecular system of primary interest is surrounded by a number of discrete molecules that constitute its environment. Various models differ by how interactions between different part of the total system are treated. The conceptually simplest approach is to do DFT calculations on the whole supersystem without singling out its individual components. Since a very large number of environment molecules often need to be added, this approach can become impractical due to steep computational scaling of conventional DFT methods. Additionally, it may be difficult to disentangle the properties of the embedded system and the environment if their states begin to mix. This can be further exacerbated by errors of the computational method, e.g., delocalization error of approximate density functionals.

These problems can be efficiently solved by employing hybrid quantum mechanics/molecular mechanics (QM/MM) methods [98, 99, 100]. The supersystem is partitioned into the embedded fragment, usually the solute molecule or the donor-acceptor complex, and the environment. The active subsystem is treated at the quantum-mechanical level (e.g., DFT), while the environment is described by a classical force field which does not explicitly account for the presence of electrons. The Hamiltonian for the total system is therefore partitioned

into the QM, MM and interaction terms

$$H_{total} = H_{QM} + H_{MM} + H_{QM/MM}, \tag{1.60}$$

where H_{QM} is the Hamiltonian accounting for interactions of nuclei and electrons of the embedded fragment, H_{MM} accounts for interaction of MM atoms, and $H_{QM/MM}$ couples these two subsystems. The wavefunction of the quantum subsystem is optimized by minimizing the energy of the total system, therefore the polarization of the embedded subsystem by the environment is taken into account through the electrostatic term in the interaction Hamiltonian. The final interaction energy is given by the expression that accounts for Coulomb interaction between solute and MM atoms represented by point charges and the additional empirical term representing van der Waals interactions:

$$E_{QM/MM} = -\int \sum_{i \in MM} \frac{q_i}{|\mathbf{R_i} - \mathbf{r'}|}\rho(\mathbf{r'})d\mathbf{r'} + \sum_{i \in MM}\sum_{j \in QM} \frac{q_i Z_j}{|\mathbf{R_i} - \mathbf{R_j}|} + \sum_{i \in MM}\sum_{j \in QM} V_{vdw}(\mathbf{R_i}, \mathbf{R_j}). \tag{1.61}$$

In simulations of CT reactions, it is important to also include the polarization of the environment due to the presence of the donor-acceptor complex. The relaxation of the environment's electron density is the major contribution to the overall outer sphere reorganization energy and is chiefly responsible for the stabilization of charged diabatic states. Including this effect in QM/MM calculations requires the use of a polarizable force field in the MM part. A simple yet effective way of including polarization of MM atoms is based on Drude oscillators [101], i.e. fictitious charged and massless particles that are harmonically bound to nuclei. Drude particles are free to move around the atom to which they are anchored in response to the electric field of other atoms and Drude particles as well the electron density of the QM region. The energy associated with Drude particles is

$$E_{Drude} = \sum_i \frac{1}{2} k_i^D |\mathbf{d_i} - \mathbf{r_i}|^2 + \sum_{ij} \frac{q_i^D q_j}{|\mathbf{d_i} - \mathbf{r_j}|} + \sum_{ij} \frac{q_i^D q_j^D}{|\mathbf{d_i} - \mathbf{d_j}|}, \tag{1.62}$$

where q_i^D and k_i^D are the charge and spring constant of the Drude particle. In practice, often a constant value is assumed for the spring constant and the charges of Drude particles are fitted to reproduce molecular polarizability or atomic polarizabilities obtained through the

distributed multipole analysis [102]. Since the position of Drude particles and the electron density of the QM region depend on each other, calculations cycle between the optimization of the wavefunction and the position of Drude charges until self-consistency is reached. In this thesis, we adopt the polarizable QM/MM method to account for the disordered and electrostatic environment for charge transfer properties of OLEDs.

1.5 Structure of this thesis

Here, we summarize the structure of the remaining parts of this thesis. In Chapter 2, we develop a new fragment-based DFT method – many-pair expansion (MPE). We start by introducing the several systematic errors in common density functional approximations, and then describe the idea of utilizing a many-pair expansion to systematically improve the accuracy of an approximate density functional. After the description of MPE formalisms, we illustrate that MPE performs very well on two important problems: (1) strong correlations in 1D/2D Hubbard and 1D Peierls-Hubbard models; (2) dispersion interactions in unsaturated hydrocarbons in the Pariser-Parr-Pople (PPP) model.

In Chapter 3, we develop a new density decomposition method – self-attractive Hartree (SAH). We first introduce the purposes for this new method: generating density fragments for applying MPE to molecular systems and extracting chemical bonding information from the electron density. After deriving the detailed algorithms for the SAH decomposition, we demonstrate that it provides an effective tool to visualize different bonding patterns from the input density. In addition, we demonstrate that it can be used to identify specific chemical bonds in molecular complexes and provides a simple and accurate electrostatic model of hydrogen bonding.

In Chapter 4, we implement MPE for molecular systems with *ab initio* Hamiltonians in Gaussian basis sets. The SAH decomposition is employed to generate localized v-representable pair densities for performing MPE calculations. We demonstrate that MPE at the second order (MPE2) already predicts accurate molecular and reaction energies for a series of small molecules and hydrogen chains, even with the low-level EXX functional as its starting point. We also show that MPE correctly describes the symmetric bond breaking in hydrogen rings, indicating its ability to remove strong correlation errors.

In Chapter 5, we start to explore the application of conventional DFT methods in com-

plex chemical systems, with the organic light-emitting diode (OLED) host-guest systems as an example. We employ a combined molecular dynamics and polarizable QM/MM method to conduct condensed phase simulations for a host-guest phosphorescent OLED. Particularly, we focus on understanding two problems: (1) how electrons and holes recombine directly on the guest emitters; (2) how how excitons are quenched and utilized for light emission. By revealing the underlying charge and energy transfer mechanisms, we suggest useful strategies for designing better OLED host-guest systems.

In Chapter 6, we study thermally activated delayed fluorescence (TADF) materials, which can be applied as efficient OLED emitters. We first describe our efforts for rational design and computational screening of TADF OLED emitters using conventional excited-state DFT methods, in collaboration with experimentalists. This study leads to the design of new highly-efficient TADF emitters and also points out the need for better computational screening protocols. Afterwards, we develop new ROKS-based screening protocols for accurate and efficient predictions of important TADF properties.

Finally, we summarize all important findings in this thesis and conclude in Chapter 7. The direction for the future work is outlined.

Chapter 2

Many-pair expansion

2.1 Introduction

Density functional theory (DFT) has been the workhorse for quantum mechanical simulations of molecules and solids due to its high performance to cost ratio [103, 104, 105]. Although there exists a universal density functional that predicts the exact ground-state electronic properties [37], its analytic form is unknown. The Kohn-Sham (KS) framework [21] is thus employed to perform DFT calculations, in which the exchange-correlation (XC) functional always needs to be approximated. Despite the great success and popularity of KS-DFT, its application can still suffer from large systematic errors because of the use of approximated XC functionals [23].

One major reason for systematic failures in density functional approximations is the self-interaction error (SIE), also known as the delocalization error, which originates from the dominating Coulomb term that pushes electrons apart [106, 107, 108]. The self-interaction error leads to underestimation of the band gaps of materials, the barriers of chemical reactions and the charge transfer excitation energies for commonly-used semi-local density functionals. Another systematic failure for approximate DFT calculations is their inability to describe degenerate or near-degenerate states, which often appear in the breaking of chemical bonds, transition metal systems and strongly correlated materials [109, 110, 111]. Such behavior occurs when the independent-particle picture breaks down, and is attributed to the strong (or static) correlation error. Furthermore, semi-local functionals cannot properly account for long-range interactions (i.e., dispersion interactions) because non-local contributions to the electron correlation are missing [112, 113], putting serious limitations on

applicability of DFT to intermolecular interactions.

Specific functionals and methods have been developed to overcome these systematic errors. For the self-interaction error, the most common solution is to add the exact exchange (EXX) into the XC functional, giving rise to hybrid [49] and range-separated [114] functionals. The challenge is to design a matching correlation functional, which would not only include the dynamic part, but also account for the static correlation. In addition, the Perdew-Zunger self-interaction correction [106] is designed to remove the spurious one-electron self-interaction in approximate functionals, though its performance is often not satisfactory [115]. Recently, the localized orbital scaling correction (LOSC) method shows promise for removing the delocalization error by properly treating the fractional electron distribution [116]. Efforts to correct strong correlation errors include adding a Hubbard-like repulsion term (DFT+U) [117], combining DFT with dynamic mean field theory [118] or multi-configurational self-consistent field calculations [119], using the strong-interaction limit of the Hohenberg-Kohn functional [120] or inverting KS potential from accurate density matrix renormalization group calculations [121]. The problem of missing dispersion interactions has also triggered development of various methods to include them in DFT calculations. The conceptually simplest approach is to add a dispersion energy correction based on the known asymptotic limit of dispersion energy, in which the leading term behaves as $-C_6/R^6$, where R is the distance between interacting atoms and C_6 coefficients are either tabulated or computed from density [122, 123, 124, 125, 126, 127, 128, 129]. A less empirical route to dispersion in DFT are van der Waals correlation functionals, which model dispersion energy through a non-local functional of the density [130, 131, 132, 133, 134]. Finally, functionals using unoccupied Kohn-Sham orbitals are also capable of accounting for long-range correlations. They either take the form of double hybrids [52, 135], which add a fraction of correlation energy calculated from perturbation theory or make use of the adiabatic connection fluctuation-dissipation theorem [136, 137, 57] to calculate the correlation functional.

As briefly overviewed, there are multiple ways to correct for a particular deficiency of an approximate exchange-correlation functional. Unfortunately, none of these is able to cure all problems at the same time. Different correction schemes often rely on empirical parameters that have been fitted to a particular data set and their reliability is difficult to predict. Also, improving one property may lead to deterioration of another as it is often the case

for self-interaction corrections [115, 138]. In other words, there is no established systematic and practical way to improve DFT results that starts from any approximation and performs uniformly better for the whole spectrum of properties and systems. Therefore, the biggest challenge in density functional development is probably to design a scheme to systematically improve the approximate functionals. One such possible DFT hierarchy is represented by Perdew's "Jacob's Ladder" [41]. According to the ingredients (density, density gradient, kinetic energy density, KS orbitals, etc.) utilized in approximate functionals, one puts the functionals onto different rungs of this hierarchy. The higher rungs are expected to give better accuracy as they use more design ingredients, but this is not necessarily true. In practice, one still needs to carefully benchmark the performance of approximate functionals for a particular class of problems before drawing any reliable conclusions.

To address this problem, we propose the many-pair expansion (MPE) [139] in this chapter, which is a hierarchy of density functional approximations that systematically corrects any deficiencies of an approximate functional to converge to the true ground-state energy. The total density is first decomposed into a sum of localized, nodeless two-electron (pair) densities (ρ_i). These pair densities are used to construct relevant four- ($\rho_i+\rho_j$), six- ($\rho_i+\rho_j+\rho_k$), ... electron densities. Numerically exact results for these few-particle densities can then be used to correct an approximate density functional via any of several many-body expansions. We first apply the proposed MPE method to several important model systems: the Hubbard [140] and Peierls-Hubbard [141] models in 1D and the Hubbard model in 2D. We show that MPE gives accurate results and is numerically convergent for strongly correlated systems: applying successively higher order corrections leads to systematic improvement of the results. To elucidate how MPE deals with long-range Coulomb interactions, we use the Pariser-Parr-Pople (PPP) [142, 143, 144] lattice Hamiltonians for polyacetylene as well as stacked ethylene and benzene molecules [145]. We show that MPE corrections decay rapidly with distance and low orders of expansion are sufficient to obtain accurate results.

2.2 Theory

2.2.1 Many-pair expansion formalism

We first describe the general formalism of the many-pair expansion (MPE) for spin-compensated systems, using the water molecule as an example, as shown in Figure 2-1. An extended for-

malism for spin-uncompensated systems (the many-electron expansion, MEE) can be found in Ref. [139]. Given a $2N$-electron system, we decompose its total electron density $\rho_T(\mathbf{r})$ into a sum of pair densities $\{\rho_i(\mathbf{r})\}$:

$$\rho_T(\mathbf{r}) = \sum_{i=1}^{N} \rho_i(\mathbf{r}), \qquad \int \rho_i(\mathbf{r}) d\mathbf{r} = 2. \tag{2.1}$$

Note that by a pair density we mean one-particle density that integrates to two electrons and not a two-particle density $\rho(\mathbf{r}, \mathbf{r}')$. For the water molecule, one obtains five pair densities from the density decomposition (1-pair terms in Figure 2-1). We will later stipulate that these pair densities need to be numerically v-representable [146]. We can further construct two-pair, three-pair, ... densities by adding several of the decomposed pair densities (2-pair and 3-pair terms in Figure 2-1). Then, assume we can compute the energy of any given electron density $\rho(\mathbf{r})$ with an approximate density functional $E_a[\rho]$ while we can obtain the exact energy $E_v[\rho]$ only for a few electrons at once. Computing the exact energy $E_v[\rho]$ is equivalent to solving the full configuration interaction (FCI) problem [147, 148] constrained to give $\rho(\mathbf{r})$ as the ground-state density. The energy correction for $\rho(\mathbf{r})$ can thus be defined as $\Delta E[\rho] \equiv E_v[\rho] - E_a[\rho]$. We then consider the following hierarchy of approximations to the true energy of this $2N$-electron system

$$\begin{aligned}
E_0[\rho_T] &\equiv E_a[\rho_T], \\
E_1[\{\rho_i\}] &\equiv E_0[\rho_T] + \sum_{i}^{N} \Delta E[\rho_i], \\
E_2[\{\rho_i\}] &\equiv E_1[\{\rho_i\}] + \sum_{i<j}^{N} (\Delta E[\rho_i + \rho_j] - \Delta E[\rho_i] - \Delta E[\rho_j]), \\
E_3[\{\rho_i\}] &\equiv E_2[\{\rho_i\}] + \sum_{i<j<k}^{N} (\Delta E[\rho_i + \rho_j + \rho_k] - \Delta E[\rho_i + \rho_j] - \Delta E[\rho_j + \rho_k] - \Delta E[\rho_i + \rho_k] \\
&\qquad + \Delta E[\rho_i] + \Delta E[\rho_j] + \Delta E[\rho_k]), \\
&\ldots
\end{aligned} \tag{2.2}$$

In this many-pair expansion, we start from the approximate DFT result ($E_0[\rho_T]$) and systematically correct its error by applying corrections calculated for a few electrons at a time. As shown in Eq. 2.2, only $E_0[\rho_T]$ is an explicit functional of $\rho_T(\mathbf{r})$, while all higher

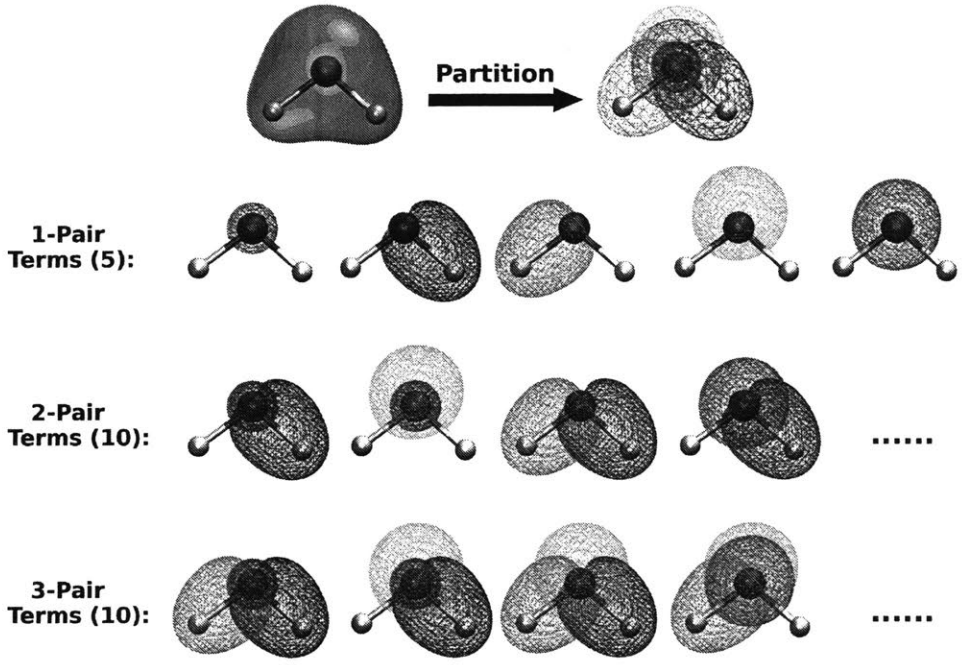


Figure 2-1: Illustration of MPE using the water molecule as an example. The pair densities are generated using the self-attractive Hartree decomposition algorithm, which will be introduced in Chapter 3.

orders are functionals of the pair densities $\{\rho_i(\mathbf{r})\}$. Once the density decomposition (Eq. 2.1) is prescribed, MPE becomes an implicit functional of the total density $\rho_T(\mathbf{r})$. This is similar to orbital-dependent functionals [149], which are implicit functionals of the density, through explicit dependence on orbitals defined by Kohn-Sham equations. We also want to point out that MPE is closely related to the many-body expansion for intermolecular interactions [150, 151], and the method of increments [152]. MPE has the important property that $E_N[\rho_T]$ gives the exact energy for $2N$ electrons no matter what approximate functional $E_a[\rho_T]$ is chosen. It thus provides a hierarchy of approximations within the context of DFT analogous to the many-body theory for the Green's function [153] and the coupled-cluster expansion of the wave function [35].

In order to compute the MPE energies, we need the approximate DFT and exact ground-state energies for various fragment densities $\rho_q(\mathbf{r})$ ($\rho_q(\mathbf{r})$ can be any one-, two-, three-, ... pair density). Assuming the approximate density functional is defined within the Kohn-Sham framework, $E_a[\rho_q]$ can be obtained via potential inversion techniques [76, 77, 78]. To

solve a non-interacting potential inversion problem for a given density $\rho_q(\mathbf{r})$, we search for a local potential $v_s(\mathbf{r})$ such that

$$\rho[v_s](\mathbf{r}) = \rho_q(\mathbf{r}), \tag{2.3}$$

i.e., the input density $\rho_q(\mathbf{r})$ is the ground-state density of a non-interacting system in the sought external potential $v_s(\mathbf{r})$. The non-interacting Slater determinant Φ is constructed out of orbitals $\{\phi_k(\mathbf{r})\}$. To find the orbitals $\{\phi_k(\mathbf{r})\}$, we search for the stationary point of the Lagrangian

$$L_{KS}[\{\phi_k\}, v_s] = \langle\Phi| - \frac{1}{2}\nabla^2|\Phi\rangle + \int v_s(\mathbf{r})\big(\rho[v_s](\mathbf{r}) - \rho_q(\mathbf{r})\big)d\mathbf{r}, \tag{2.4}$$

where $v_s(\mathbf{r})$ is the Lagrange multiplier enforcing the density constraint. For a given $v_s(\mathbf{r})$, each KS orbital $\phi_k(\mathbf{r})$ satisfies a one-electron Schrödinger equation:

$$-\frac{1}{2}\nabla^2\phi_k(\mathbf{r}) + v_s(\mathbf{r})\phi_k(\mathbf{r}) = \epsilon_k\phi_k(\mathbf{r}), \tag{2.5}$$

and the density can then be expressed as

$$\rho[v_s](\mathbf{r}) = 2\sum_{k}^{occ}\phi_k(\mathbf{r})^2. \tag{2.6}$$

Since the potential $v_s(\mathbf{r})$ that yields $\rho_q(\mathbf{r})$ through Eqs. 2.5 and 2.6 is not known, we need to find it numerically to satisfy Eq. 2.3. We start with an initial guess $v_s^0(\mathbf{r})$ and use Newton's method [154] to find an improved potential

$$v_s^{j+1}(\mathbf{r}) = v_s^j(\mathbf{r}) + \int \left(\frac{\delta\rho(\mathbf{r}')}{\delta v_s(\mathbf{r})}\bigg|_{v_s^j}\right)^{-1}\left(\rho_q(\mathbf{r}') - \rho[v_s^j](\mathbf{r}')\right)d\mathbf{r}', \tag{2.7}$$

where $\delta\rho(\mathbf{r}')/\delta v_s(\mathbf{r})$ denotes the response kernel, which can be computed analytically from the perturbation theory

$$\frac{\delta\rho(\mathbf{r}')}{\delta v_s(\mathbf{r})} = 4\sum_{i}^{occ}\sum_{a}^{vir}\frac{\phi_i(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r}')}{\epsilon_i - \epsilon_a}. \tag{2.8}$$

The Newton's search using Eqs. 2.5-2.8 is iterated until Eq. 2.3 is satisfied. Then the DFT energy for $\rho_q(\mathbf{r})$ can be computed using the converged KS orbitals: $E_a[\rho_q] = E_a[\{\phi_k\}]$.

To obtain the exact energy of the same density $\rho_q(\mathbf{r})$, we can apply a similar potential

inversion construction except that now we are searching for the exact wave function in a fully interacting system [155]. The interacting wave function Ψ is a stationary point of the Lagrangian

$$L_{Exact}[\Psi, v_{ex}] = \langle\Psi|[\frac{1}{2}\sum_k \hat{\mathbf{p}}_k^2 + \sum_{k<l}\frac{1}{\hat{\mathbf{r}}_{kl}}]|\Psi\rangle + \int v_{ex}(\mathbf{r})\big(\rho[v_{ex}](\mathbf{r}) - \rho_q(\mathbf{r})\big)d\mathbf{r}, \tag{2.9}$$

where $v_{ex}(\mathbf{r})$ enforces the density constraint. The interacting wave function Ψ is the ground state of the interacting Schrödinger equation

$$[\frac{1}{2}\sum_i \hat{\mathbf{p}}_i^2 + \sum_{i<l}\frac{1}{\hat{\mathbf{r}}_{il}} + v_{ex}(\hat{\mathbf{r}})]|\Psi\rangle = E|\Psi\rangle, \tag{2.10}$$

and the corresponding density is

$$\rho[v_{ex}](\mathbf{r}) = \langle\Psi|\delta(\hat{\mathbf{r}} - \mathbf{r})|\Psi\rangle. \tag{2.11}$$

The search for the external potential $v_{ex}(\mathbf{r})$ employs the similar algorithm as in the non-interacting case, except that we use numerical finite differences to compute the response kernel $\delta\rho(\mathbf{r}')/\delta v_{ex}(\mathbf{r})$. Once the converged potential $v_{ex}(\mathbf{r})$ is found, the exact energy can be obtained as the expectation value of the interacting Hamiltonian $E_v[\rho_q] = \langle\Psi|\hat{\mathbf{H}}|\Psi\rangle$.

From the process of computing $E_a[\rho_q]$ and $E_v[\rho_q]$, it is clear that $\rho_q(\mathbf{r})$ must be the ground-state density of some potential so that we can find its corresponding wave function and energy. In other words, the fragment density $\rho_q(\mathbf{r})$ must be non-interacting and interacting v-representable, which needs to be enforced by the density decomposition in Eq. 2.1. Although the sufficient conditions for v-representability are not known [156], this is unlikely to be a problem in practice provided that pair densities are sufficiently smooth. Note that even non-v-representable densities can be approached arbitrarily close with smooth potentials [157, 158] and that within the finite basis set approximation even node-containing densities can be effectively v-representable [159].

2.2.2 Relation to Perdew-Zunger self-interaction correction

The MPE (or MEE) hierarchy systematically removes any possible errors of an approximate density functional $E_a[\rho]$. In particular, it can be seen as a method to systematically remove the self-interaction error, which is one of the major sources of density functional approx-

imation failures. For this reason it is instructive to establish a connection between MPE and the Perdew-Zunger self-interaction correction (PZ-SIC) [106]. PZ-SIC is exact for any one-electron system, while MPE converges to the exact energy for an arbitrary number of electrons; therefore, the latter can be viewed as a generalization of the former. Nevertheless, several important differences need to be noted.

For an N-electron system, the PZ-SIC can be written in the following form

$$E_{SIC}[\{\rho_i\}] = E_a[\rho] + \sum_i^N \left(E_{EXX}[\rho_i] - E_a[\rho_i]\right), \tag{2.12}$$

where E_a is an approximate functional of the total energy, E_{EXX} is the functional using the exact exchange for the exchange-correlation part and the sum runs over a set of one-electron densities. For a one-electron system ($N = 1$), PZ-SIC is equivalent to the first order of MEE (which uses one-electron density as the smallest fragment), because in the absence of electron-electron interactions E_{EXX} is the exact functional. In a many-electron case, Eq. 2.12 and MEE1 are still formally equivalent, but they are defined on a different set of admissible one-electron densities $\{\rho_i\}$. In PZ-SIC $\rho_i = |\phi_i|^2$ are orbital densities, where $\{\phi_i\}$ are occupied Kohn-Sham orbitals or localized orbitals obtained through a unitary transformation. As orbitals are in general excited states of a one-electron Schrödinger equation, they exhibit pronounced nodal surfaces and, as a result, the corresponding densities are not smooth. Considering that $E_a[\rho]$ is an approximation to the ground-state functional, its performance for such excited-like, non-smooth densities is likely to be much worse than for the total density, which naturally is a ground state of some potential. This caveat has been linked to the often unsatisfactory performance of PZ-SIC [115]. The complicated nodal structure of higher-lying orbitals is a consequence of their mutual orthogonality. Recently, it has been proposed to use complex orbitals [160, 161, 162] in PZ-SIC as the nodes of the real and imaginary parts do not need to coincide, so the resulting densities are considerably smoother. A similar problem with orbital nodes appears for functionals using the ratio of von Weizsäcker and exact non-interacting kinetic energy densities to detect one-electron regions of the density [163, 164]. Recently a density-based alternative to this detector has been shown to be free of the nodal problem [165, 166, 167]. In MEE one-electron densities $\{\rho_i\}$ are required to be v-representable (be ground states of some potential). In practice, this means that orbital densities are not allowed and that admissible densities should be smooth.

This requirement makes the energies $E_a[\rho_i]$ compatible with the approximate total energy $E_a[\rho]$, as in both cases a ground-state functional is applied to densities that are known to be ground states. V-representability of one-electron densities needs to be assured by the initial decomposition, which is not uniquely defined. This is a similar situation as for PZ-SIC, which is not invariant with respect to orbital rotations. In both cases, the partitioning of the total density may be optimized variationally.

Another possible reason why PZ-SIC can deteriorate some results is that the SIE often mimics to some extent static correlation [168], so removing it destroys the helpful compensation of errors in DFT. PZ-SIC is designed to remove only the one-electron SIE, which is easy to define for every orbital separately and can be eliminated by the use of the one-electron self-interaction-free exact exchange functional. Unfortunately, as PZ-SIC does not add any correlation beyond the initial approximation, it can destroy the balance between exchange and correlation functionals in modeling of the total exchange-correlation hole. Also, PZ-SIC has no means to correct for the many-electron self-interaction error [169], which is also pervasive in approximate density functionals. In contrary, MPE (MEE) is able to systematically remove the many-electron SIE and include the corresponding degree of many-electron correlation. In the MPE formalism pair densities are treated at the first order, so MPE1 in addition to correcting for the one-electron self-interaction corrects the correlation energy of two electrons within each pair. This allows to account for possible static correlation and for example correctly dissociate the H_2 molecule. At higher levels, more electrons are correlated at a time, so many-electron self-interaction and many-electron strong correlations are treated in a balanced way.

2.3 Results: strong correlation

2.3.1 One-dimensional Hubbard model

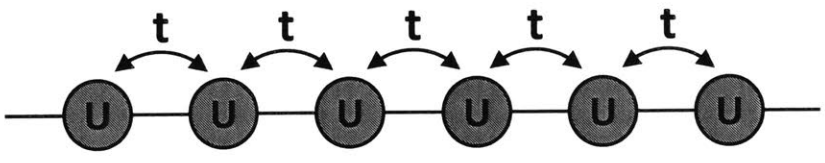


Figure 2-2: Illustration of one-dimensional Hubbard model.

To test the performance of MPE on strongly correlated systems, we first consider the one-dimensional Hubbard model (shown in Figure 2-2) described by the Hamiltonian

$$\hat{H} = \sum_{i\sigma} t_i \left(\hat{a}^\dagger_{i,\sigma} \hat{a}_{i+1,\sigma} + \hat{a}^\dagger_{i+1,\sigma} \hat{a}_{i,\sigma} \right) + U \sum_i \hat{a}^\dagger_{i,\alpha} \hat{a}_{i,\alpha} \hat{a}^\dagger_{i,\beta} \hat{a}_{i,\beta}, \tag{2.13}$$

where $t_i \equiv t$. The first term describes hopping of electrons between neighboring sites and the second describes on-site repulsion of opposite spin electrons characterized by the parameter U. This model has been used extensively in solid-state physics to describe the transition between conducting and insulating systems, and also novel superconducting effects. It describes potentially strongly correlated electrons on a lattice and often serves as a benchmark for electronic structure methods [170, 171] as the exact solution, based on the Bethe Ansatz, is known [172]. Different formulations of DFT exist for the Hubbard model, which differ by the choice of the basic variable in lieu of real-space density [173, 174, 175]. In this work, the density of the system is understood as the diagonal of the density matrix in the site basis. Since the model has translational symmetry under periodic boundary conditions (p.b.c.), the ground-state density is equal at each site and amounts to $\rho_\alpha = \frac{2N_{occ}}{N}, \forall \alpha = 1, \ldots, N$, where N_{occ} is the the number of electron pairs distributed over N sites of the lattice.

As the total ground-state density is known, we can decompose it to a sum of pair densities (Eq. 2.1), in principle, by any prescription that assures v-representability. One of the necessary conditions for a density to be v-representable on a lattice with p.b.c. is that it is positive at each site [156]. Here, we relax this restriction, requiring only that ρ_i be non-negative, which is feasible if we also allow infinite v_i . In practice, we want to partition total density in such a way that pair densities are compact. Physically, this would allow to interpret them as localized electron pairs and to capture most of correlation. A viable procedure to achieve such decomposition is to recall the Boys orbital localization criterion [176], which minimizes the spatial spread of orbitals. Applying this procedure to the 1D Hubbard model results in pair densities composed of contiguous blocks (Figure 2-3). All numerical results presented in this section are based on this prescription.

For a typical filling, the localization results in pair densities that are inhomogeneous: they equal N_{occ}/N for the central site(s) but only contain part of the density on the edge sites (see Figure 2-3a). In this case, the MPE1 correction is a sum of many slightly different

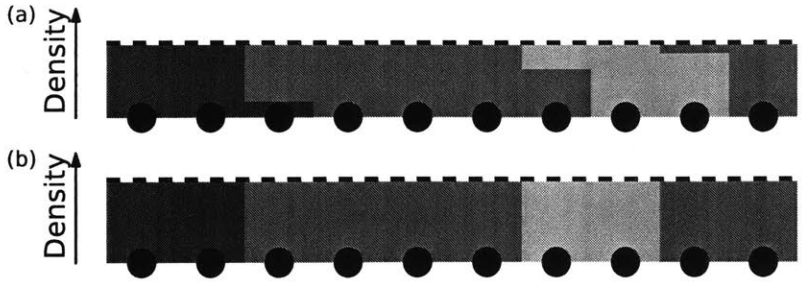


Figure 2-3: Localized pair densities for 1D Hubbard model at (a) non-periodic filling ($\langle n \rangle$=0.9), (b) periodic filling ($\langle n \rangle$=1.0). Lattice sites are represented by black circles, while each pair density is marked by one color. The black dashed line is the total density for each site. Note that the pair density picture is incomplete in (a) because remainder of the lattice is truncated.

energies:

$$E_1[\rho] = E_0[\rho] + \Delta E[\rho_1] + \Delta E[\rho_2] + \Delta E[\rho_3] + ... \tag{2.14}$$

However, for certain fillings (such as 1/2 or 1/3), the partitioning procedure leads to pair densities that repeat periodically along the chain (see Figure 2-3b), resulting in an energy $E_1[\rho] = E_0[\rho] + N_{occ}\Delta E[\rho_1]$. Note that this transition happens abruptly - starting from a periodic filling and adding even one electron pair results in a completely aperiodic filling. As a result, the MPEn energy is not a smooth function of filling. A way to resolve this is to average energies over different possible pair density partitions. In practice, we perform the averaging by adding an additional constraint on $\rho_{1,1} = \gamma$ and integrating over $\gamma \in (0, \langle n \rangle]$

$$E_{AMPE}[\rho] = \int_0^{\langle n \rangle} E_{MPE}(\gamma) d\gamma \tag{2.15}$$

We evaluate this integral by quadrature, which then directly mimics the average that is done for the aperiodic filling case (Eq. 2.14).

As the approximate functional in Eq. 2.2, we use the exact exchange (EXX) and local density approximation (LDA). Exact diagonalization (full configuration interaction) is used to compute $E_v[\rho]$. Because LDA can be exact for homogeneous Hubbard model by design, we fit the LDA energy per site to the Bethe Ansatz energies to obtain LDA exchange-correlation energies. Here we show the fitted LDA exchange-correlation energies in Figure 2-4, which we use for our LDA-(A)MPE calculations.

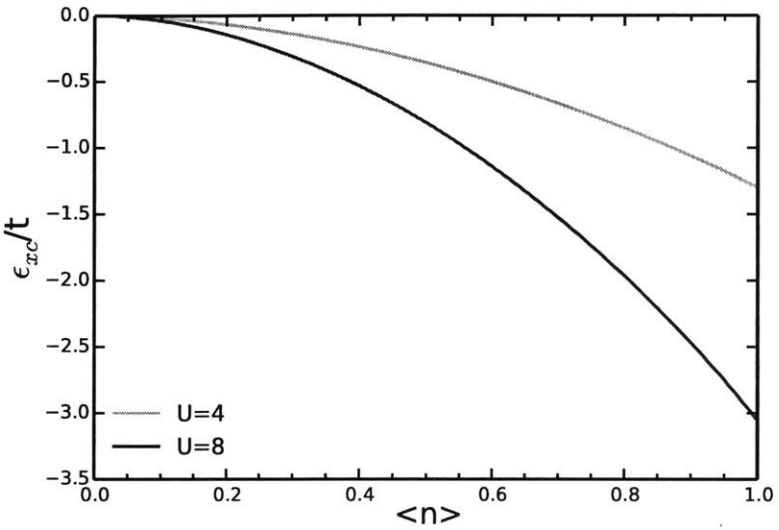


Figure 2-4: Fitted LDA exchange-correlation energy per site as a function of site occupancy $\langle n \rangle$.

To compute EXX (LDA) and exact energies for the fragment density ρ_q, we need to search for the potentials v_s and v_{ex} in Eq. 2.5 and Eq. 2.10. When doing the potential inversion calculations (Eqs. 2.3-2.11) in the site basis, the density and potential are represented as vectors $\boldsymbol{\rho}$ and $\boldsymbol{v}$, while the coordinates $\mathbf{r}$ and $\mathbf{r}'$ become the lattice site indices. As a result, $\delta\rho(\mathbf{r})/\delta v(\mathbf{r}')$ is replaced by a Jacobian matrix $d\boldsymbol{\rho}/d\boldsymbol{v}$. We start the calculations by building the non-interacting and interacting N_s-site Hubbard Hamiltonians with the initial guess for potential $\boldsymbol{v}_0(1, 2, ..., N_s)$ such that for site α, $\boldsymbol{v}_0(\alpha) = \infty$ if $\boldsymbol{\rho}_q(\alpha) = 0$. By diagonalizing the Hamiltonian $\mathbf{H}[\boldsymbol{v_0}]$, we obtain the eigenvalues $\{\epsilon_i\}$ and eigenvectors $\{\boldsymbol{\phi}_i\}$ and the corresponding density $\boldsymbol{\rho}[\boldsymbol{v}_0]$. As the Jacobian $d\boldsymbol{\rho}/d\boldsymbol{v}$ may be ill-conditioned or singular, to compute Newton steps in Eq. 2.7 we use Tikhonov regularization [154]

$$[\boldsymbol{v}_{i+1} - \boldsymbol{v}_i] = \boldsymbol{V}\boldsymbol{D}\boldsymbol{U}^{\boldsymbol{T}}[\boldsymbol{\rho}_q - \boldsymbol{\rho}[\boldsymbol{v}_i]] \tag{2.16}$$

$$\boldsymbol{D}_{ii} = \frac{\sigma_i}{\sigma_i^2 + \lambda}, \tag{2.17}$$

where σ_i are singular values of the Jacobian and $\boldsymbol{U}, \boldsymbol{V}$ are unitary matrices obtained through singular value decomposition (SVD) [86, 177]. We set the regularization parameter $\lambda = 10^{-6}$.

In Figure 2-5, we plot the averaged MPE (AMPE) energy curve of a 500-site 1D Hubbard model as a function of the site occupancy $\langle n \rangle$. We perform MPE calculations up to the 4^{th}

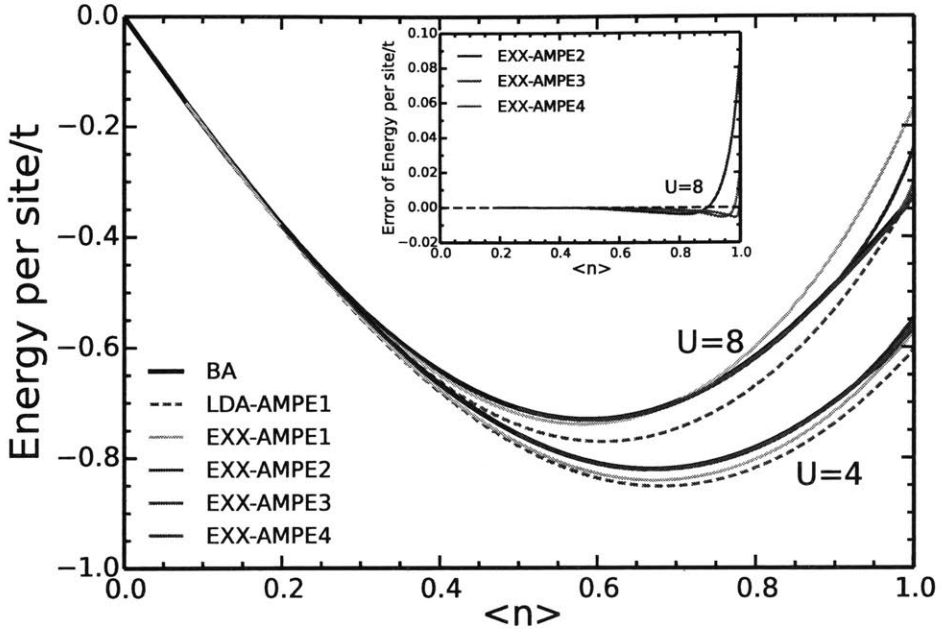


Figure 2-5: Energy per site and its errors for 1D Hubbard model as a function of site occupancy $\langle n \rangle$.

order (Eq. 2.2), which means we only need to do exact calculations on up to 4 electron pairs at a time. Thanks to the locality of interactions (A)MPE at any level scales linearly with the system size as opposed to factorial scaling of exact diagonalization (FCI). The exact Bethe Ansatz (BA) results are presented for comparison. Overall, the AMPE energy curves are in excellent agreement with the BA curve. Even at 1^{st} order, EXX-AMPE is in good agreement with the reference, whereas LDA-AMPE deviates more significantly. Considering that, by design, LDA is exact for the homogeneous Hubbard model, the poor performance of LDA-AMPE1, teaches us something about LDA: while it is exact for the uniform system, treatment of two- and many-electron interactions is unbalanced. Adding in the correct interactions for each pair then makes the results worse because the many-electron errors are exposed and only summation up to the N-pair contribution makes the resulting errors cancel. Starting from the 2^{nd} order, curves representing EXX-AMPE and LDA-AMPE energies become visually indistinguishable and the latter are suppressed in Figure 2-5 for clarity. As can be seen, when we apply successive higher order corrections, the AMPE energies converge quickly towards the exact result, which confirms that our method can be systematically improved. For reference, the MPEn energies are visually indistinguishable

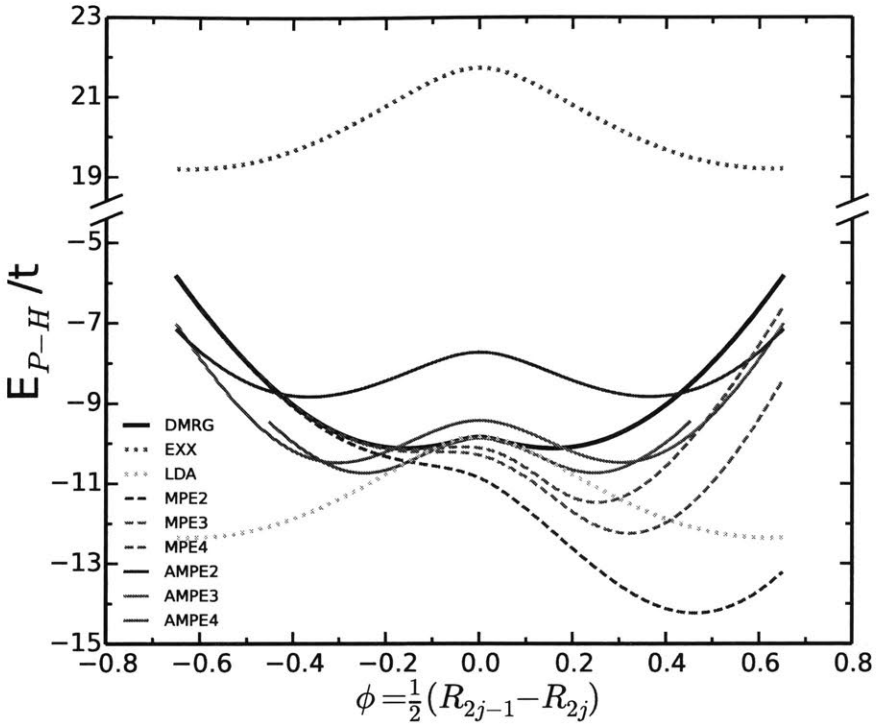


Figure 2-6: Energy for 30-site Peierls-Hubbard model as a function of bond shift parameter ϕ (U=8, $\langle n \rangle$=1, $\omega = 2.8t$).

from the AMPEn energies, except at the periodic fillings, where the MPEn results would be discontinuous.

2.3.2 One-dimensional Peierls-Hubbard model

As the second example, consider the Hubbard model with a Peierls distortion ($t_{2j-1} \neq t_{2j}, \quad j = 1, \ldots, N/2$)[141]. This model reflects the spontaneous symmetry breaking of the 1D periodic lattice resulting, for example, from alternating single and double bonds in a conjugated polymer like polyacetylene. Such displacement with respect to the symmetric Hubbard model can be described with a bond shift parameter $\phi = \frac{1}{2}(R_{2j-1} - R_{2j})$, where R_i denotes the bond length between sites i and $i+1$. The Hamiltonian takes the form of Eq. 2.13 with $t_{2j-1} = te^{-\phi}$ and $t_{2j} = te^{\phi}$. Keeping the analogy with polyacetylene, we recognize that the Peierls-Hubbard model only treats the π electrons. To incorporate the additional energy cost of stretching and squeezing the underlying σ bonds, for a given displacement ϕ we add a harmonic term $N\frac{\omega}{2}\phi^2$ to the total energy, where ω=2.8t gives approximately the correct

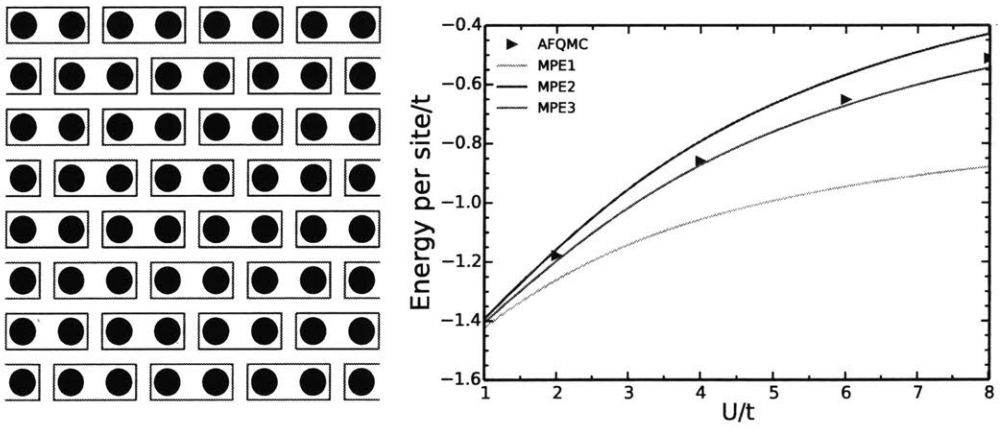
physics for U=8. We restrict our model to 30 sites, in which case numerically exact results are easily obtained from DMRG [178, 179]. As can be seen in Figure 2-6, the exact curve has a symmetric double well shape characteristic of the expected symmetry breaking. EXX results reproduce this qualitative feature, but are far too high in energy and the predicted bond shifts are too large. LDA results in a uniform downward shift with respect to EXX such that the LDA energy is correct for $\phi = 0$. For $\phi \neq 0$, LDA predicts energies which are far too low, which is again a manifestation of many-electron self-interaction errors in LDA.

At half filling, the Peierls-Hubbard model presents an interesting challenge for MPE. Assuming partitioning into non-overlapping pair densities (Figure 2-3b), EXX-MPE and LDA-MPE are equivalent as the LDA correction is exactly canceled out at the 1^{st} order. Averaging leads only to a uniform shift, therefore, only EXX-(A)MPE is explicitly considered in Figure 2-6 and in the following discussion. The pair densities (Figure 2-3b) naturally break the symmetry of the lattice when $\phi \neq 0$ - the pairs either localize on a $2j-1, 2j$ bond or on a $2j, 2j+1$ bond. In the former case, for $\phi < 0$ ($t_{2j-1} > t_{2j}$), the short double bonds are located between two sites occupied by the same density pair, MPE therefore will give a better description than for $\phi > 0$, where the short double bonds are located between different density pairs. This is clearly demonstrated in Figure 2-6, where the $2j-1, 2j$ is chosen, leading to a very accurate treatment for negative ϕ, but strong overcorrelation for positive ϕ. Particularly at high orders, MPEn does an impressive job of reproducing the energy dispersion about the minimum, but the global behavior is unsatisfactory.

To recover the symmetric shape, we again apply the averaging procedure for MPE. By averaging over different pair density partitions, AMPE results do not rely on particular pair density positions. Thus, AMPE avoids MPE's asymmetry problem and finds two local minima correctly. The AMPEn minima clearly approach the DMRG ones as n increases. Nevertheless the convergence to the exact result is rather slow. On the other hand, when $\phi < 0$, MPE is more accurate than AMPE, which suggests that some a priori knowledge of the electronic structure could perhaps be used to improve the results - an ansatz capable of picking out the "best" density pattern might be able to capture MPE's accuracy near the minimum together with AMPE's global symmetry.

2.3.3 Two-dimensional Hubbard model

Finally, we note that MPE is not in any way restricted to 1D systems. For example, Figure 2-7 show results for the 2D Hubbard model, whose sites form a two-dimensional square lattice. Due to the macroscopic degeneracy of the model, there are many equivalent density partitionings making it difficult to arrive at definitive MPEn numbers for the model. Still, as an illustration we present MPEn energy curves for 8×8 sites at half filling (Figure 2-7), using an aligned $A - A - A - A$ arrangement of the 1D partitions in Figure 2-3b to tile 2D space. We use auxiliary field quantum Monte Carlo (AFQMC) [180] data as reference. It can be seen in the figure that MPE gives accurate results for half-filling 2D Hubbard model, even at U=8. Similarly to the 1D case, MPE results are also systematically improved by applying higher order corrections in the 2D model. For instance, at U=4, the EXX-MPEn error relative to the AFQMC estimates goes from 23% to 7.8% to 1.5% as n goes from 1 to 3. This result demonstrates the applicability of MPE to higher dimensional systems.



(a) (b)

Figure 2-7: Performance of MPE for 2D Hubbard model. (a) Pair densities for 2D Hubbard model at half filling. Each density pair is marked by a blue rectangle. Note that the pair density locations are shifted by one site every other row. Periodic boundary conditions are applied in both directions. (b) Energy per site for a 2D Hubbard model of 8×8 sites at half filling as a function of correlation strength U.

2.4 Results: dispersion interaction

Dispersion forces result from interaction of instantaneous dipole moments, which in turn are generated by electron correlations. This is a weak but long-range effect, which lies at the

heart of supramolecular chemistry and is decisive for the existence and structure of many systems of biological or technological importance. Unfortunately, being an effect of electron correlations, dispersion interactions are completely absent in the Hartree-Fock theory. Also, while the exact exchange-correlation functional takes account for these interactions, common semi-local approximations to the correlation functional are intrinsically unable to capture this effect due to its fundamentally non-local character. In this section, we discuss the performance of MPE on dispersion interactions, using lattice models for unsaturated hydrocarbons as an example.

2.4.1 Pariser-Parr-Pople model

To model the electronic structure of π electrons, we employ the Pariser-Parr-Pople (PPP) model, defined by following Hamiltonian

$$\hat{H} = \sum_{i\sigma} t_{ij} \left(\hat{a}^\dagger_{i,\sigma} \hat{a}_{j,\sigma} + \hat{a}^\dagger_{j,\sigma} \hat{a}_{i,\sigma} \right) + U \sum_i \hat{a}^\dagger_{i,\alpha} \hat{a}_{i,\alpha} \hat{a}^\dagger_{i,\beta} \hat{a}_{i,\beta} + \frac{1}{2} \sum_{i \neq j} V_{ij} (\hat{a}^\dagger_i \hat{a}_i - 1)(\hat{a}^\dagger_j \hat{a}_j - 1). \quad (2.18)$$

This can be viewed as an extension of the Hubbard model, which in addition to on-site repulsion characterized by the parameter U, accounts for long-range electron-electron interactions determined by the inter-site potential V_{ij}. A common choice for V_{ij} is the Ohno potential [181], which is an interpolation between the on-site interaction U and asymptotic $1/r$ dependence of the Coulomb potential. The Ohno potential has the following form

$$V_{ij} = \frac{U}{\sqrt{1 + (U r_{ij}/e^2)^2}}, \quad (2.19)$$

where r_{ij} is the distance between sites i and j, and e is the elementary charge. The parameters t_i and U are usually fitted to reproduce experimental data. For two carbon atoms separated by 1.4 Å, the standard PPP parameters are $U = 11.26$ eV and $t_{ij} = -2.4$ eV. Further, we assume exponential dependence of the transfer integral on the bond distortion $\phi = (1.4 - r_{ij})$, i.e. $t_{ij} = -2.4e^{\alpha\phi}$ eV, for covalently bound sites. The parameter $\alpha = 3.785$ was fitted to reproduce the transfer integral $t = -2.9$ eV for ethylene ($r = 1.35$ Å).

Further in this work we consider only hydrocarbons whose all site occupations are equivalent by symmetry. Additionally, each sp^2 carbon atom contributes only one electron to the π framework, so we will consider only half-filled PPP models. This means that for each site

k the ground-state density is just $\rho_k = 1$. We impose these constraints to avoid the need for optimization of the density. This restriction is not a limitation of MPE itself, but makes the initial study simpler and faster without any substantial loss of generality.

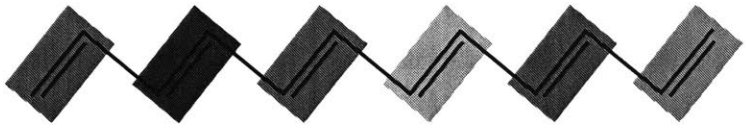


Figure 2-8: Density partitioning in a half-filled system with bond-alternation. Each color block represents one pair density.

For the models considered, the density partitioning is particularly straightforward as it is always possible to choose a pattern where each pair occupies two neighboring sites and all pairs do not overlap with each other. For systems with alternating bond lengths, we allocate a pair to sites connected by a shorter bond, while longer bonds connect sites occupied by two different pairs (Figure 2-8). We have shown that such decomposition leads to a quick convergence of MPE energies for the Peierls-Hubbard model. This also corresponds well with intuition that tells us that electron pairs would localize on double (short) carbon-carbon bonds rather than on single (long) ones.

2.4.2 Decay of MPE corrections in polyacetylene

Due to locality of interactions in the Hubbard model, MPE has $O(N_s)$ scaling at any order, where N_s is the size of the system. In the PPP model, the presence of long-range interactions, requires to calculate energy corrections for every possible combination of pair densities. This immediately leads to $O(N_s^M)$ where M is the order at which MPE is truncated. However, we can expect that energy corrections will decay rapidly with the distance between electron pairs that are correlated. If we could neglect contributions from combinations of pair densities separated by a distance greater than some threshold, MPE would still scale linearly with the size.

To check how rapidly MPE correction vanish with the distance between pair densities, we first consider a 100-site model for trans-polyacetylene with imposed cyclic boundary conditions. The bond lengths are 1.36 Å and 1.44 Å for the the double and single bond, respectively. We consider four types of pair density patterns (see Figure 2-9) that have one gap in the partitioning. For MPE2, one pair density is fixed and the other one is moved along the chain (MPE2 (1+1)). Similarly, for MPE3 (2+1) and MPE4 (3+1), two and three

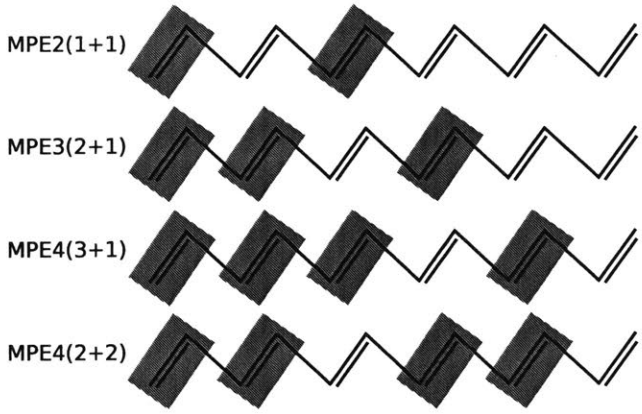
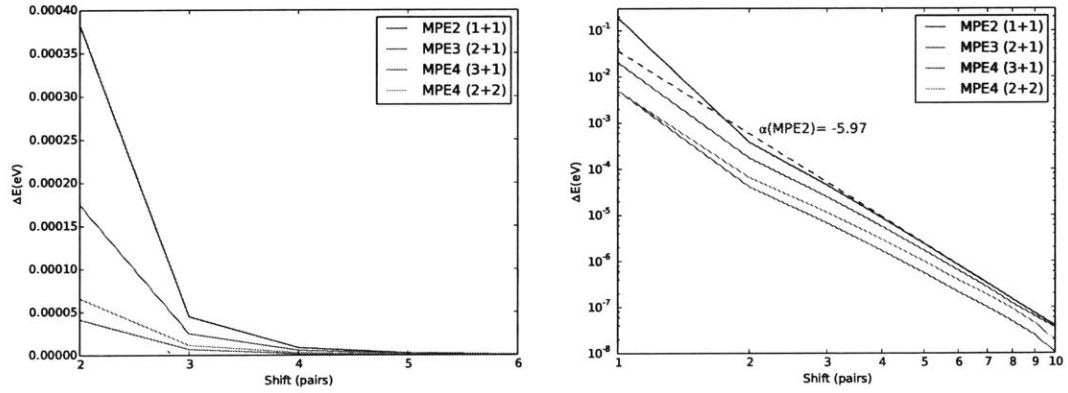


Figure 2-9: Four pair density patterns considered in the analysis of the MPE corrections decay rate. Each block represents one pair density and the blocks with the same color are fixed as contiguous. Shift = 2 for all patterns.

adjacent pair densities are fixed, and the remaining one is translated by two sites at a time. Additionally, for MPE4 we consider a (2+2) pattern, where two adjacent pair densities are fixed and the remaining two are translated as contiguous blocks.



(a) (b)

Figure 2-10: (a) Decay of MPE corrections as a function of the distance between pair densities in polyacetylene. Shift denotes the distance between two contiguous blocks of fragment densities measured in number of pair density translations. (b) Decay of MPE corrections on a double logarithmic scale and a monomial fitted to be tangent to the MPE2 curve at shift=10.

Figure 2-10a shows decay of the magnitude of MPE2 - MPE4 energy corrections as a function of separation between localized pair densities. Energy corrections for contiguous blocks of pair densities (shift = 1) are suppressed as they are up to three orders of magnitude

larger than for a pattern with a gap between the pairs. It is evident that corrections decay rapidly and fall below 10^{-6} eV beyond the separation of 5 localized pairs (10 sites). More insight can be gained from plotting the same data on a double logarithmic scale (Figure 2-10b), this time including also the corrections for contiguous blocks. For shift = 1, the corrections are 0.19 eV, 0.020 eV and 0.0050 eV for MPE2, MPE3 and MPE4, respectively. This already shows that, if pairs are localized in space, corrections decay quickly with the level of the MPE, each being an order of magnitude smaller than the preceding one. Furthermore, the magnitudes of corrections fall off rapidly also with the separation between localized pairs. The decay is polynomial as the dependence on the separation is approximately a straight line on a double logarithmic scale. The blue broken line in Figure 2-10b is tangent to the MPE2 curve at shift = 10 and it follows closely the MPE2 curve for shifts greater than 4. The slope is -5.97, which means that MPE2 corrections decay as $1/R^6$ in the asymptotic range, which corresponds to the fact that they account for unscreened long-range correlations due to dispersion interactions.

The rapid decay of MPE corrections, both with the level of the expansion and the distance between pairs being correlated, is very promising for the applications of MPE. First, we can truncate the expansion at low orders as there is no need to correlate more than a few electron at a time. This behavior has already been shown for Hubbard and Peierls-Hubbard Hamiltonians. Second, we can suppress the corrections for pairs separated by large distances, which would reduce the polynomial $O(N_s^M)$ scaling at the cost of some small loss in accuracy. Both these features hinge on the decomposition of the total density to well localized pair densities, which is an essential component of MPE calculations.

2.4.3 Ethylene dimer

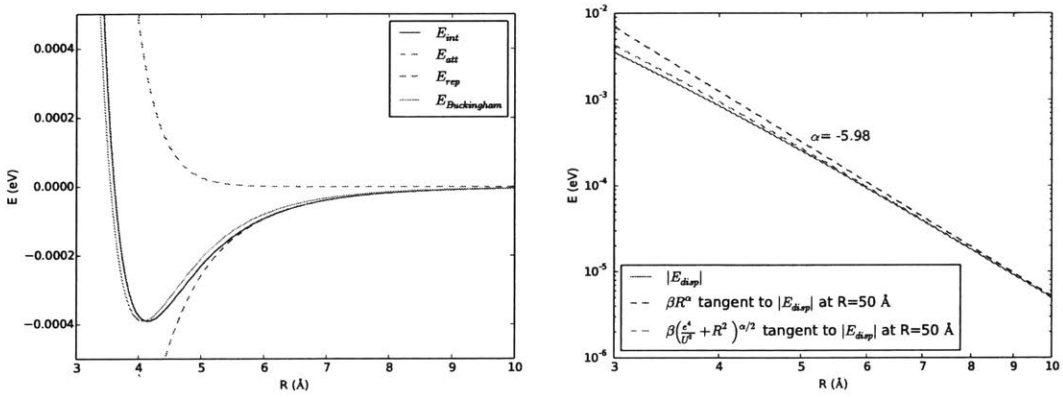
To study how MPE accounts for dispersion interactions, we first start with a model for the ethylene dimer. Dispersion interactions in this system are dominated by correlations of π electrons, which are the most polarizable. In reality different orientations of the dimer strongly affect the interaction energy [182]. This includes also the dispersion contribution due to the directional character of the π cloud. The anisotropy cannot be accounted for in the model Hamiltonian that we use, which is represented in the basis of four sites in the positions of carbon atoms only. Therefore, we stipulate that our model represents a configuration where two ethylene molecules are stacked perfectly on top of each other and

interact mostly through $\pi - \pi$ stacking.

Since there are no electron clouds that could overlap in the lattice version of the PPP model and all sites are charge neutral, there is no interaction between two molecules at the mean-field level. Allowing electrons on both molecules to correlate, generates a force that is attractive at all separations. The lack of short-range repulsion due to wavefunction overlap causes that MPE interaction energies do not behave like in real π-π interacting complexes. In particular, they cannot predict the existence of an equilibrium distance. To account for this effect, we add a pairwise correction to the energy that depends exponentially on distance between interacting sites:

$$E_{rep} = \sum_{i \in A, j \in B} \epsilon \exp(-r_{ij}/D) \tag{2.20}$$

where $\epsilon = 377.2$ eV, $D = 0.3455$ Å, and r_{ij} is the distance between sites i and j belonging to two different molecules. The parameters of this potential have been chosen so that the total MPE2 interaction energy closely follows the Buckingham potential from the MM3 force field [183] at all distances. Note that when comparing energy differences at fixed distances, the repulsive part will cancel out, so its particular form is somewhat arbitrary in this case.



(a) (b)

Figure 2-11: (a) Dissociation curve of the ethylene dimer calculated from MPE2 with a repulsive correction and the decomposition to its attractive and repulsive components. For comparison, a Buckingham potential from the MM3 force field using the same parameters is shown. (b) Dispersion energy of the ethylene dimer calculated from MPE2 and monomial fits to its asymptotic limit (double logarithmic scale).

Figure 2-11a shows the dissociation curve calculated at the MPE2 level, which is exact for the four-electron PPP Hamiltonian used to model the interaction. Naturally, the curve

has the qualitative character of the potential energy surfaces of a van der Waals complex. The shapes of MPE2 and Buckingham potentials are very similar at all separations, differing mostly by a slight shift at the short and intermediate range, which justifies the parameterization of the repulsion correction we use. The attractive (MPE2) and repulsive components of the total interaction energy are also plotted in Figure 2-11a. To analyze the behavior of the attractive part of the total MPE2 (exact) interaction energy, we plot its absolute value on the double logarithmic scale (Figure 2-11b). Since the attractive component of the MPE2 interaction energy is pure dispersion, it should scale like R^{-6} at large separations between monomers. The broken blue straight line in Figure 2-11b is a monomial fitted to be tangent to the dispersion energy curve at $R = 50$ Å. Its slope is -5.98, which matches the asymptotic limit almost perfectly. The exact dispersion energy approaches this limit rather slowly and at 10 Å the slope is still noticeably smaller. This is expected since the Ohno potential that has been used in the PPP Hamiltonian reaches the $1/r$ dependence of the true Coulomb potential only at infinity. For this type of two-electron interactions, the correct long-range scaling of the dispersion energy is $\left((e^4/U^2) + R^2\right)^{-3}$ (see Eq. 2.19). The red curve in Figure 2-11b is a similarly fitted monomial of $\sqrt{(e^4/U^2) + R^2}$ plotted as a function of R. It follows the exact dispersion energy much closer and around 7 Å these two curves become visually indistinguishable.

2.4.4 Benzene dimer

As a second example, we study the model for a parallel stacked benzene dimer. The same parameters have been used as for the ethylene dimer, both in the PPP Hamiltonian and the repulsive energy correction. Assuming that the density is decomposed into maximally localized pairs, two types of partitioning are possible (Figure 2-12). Unless otherwise stated, we will assume the partitioning of the D_{3h} symmetry, which makes both benzene molecules symmetric. Figure 2-13 shows MPE potential energy surfaces up to the 6th order, which is equivalent to the exact diagonalization (FCI) for the entire system.

As the MPE1 energy adds only a portion of intramolecular correlation energy, the interaction energy is exactly the same as for the EXX, so both curves are on top of each other. MPE2 adds pairwise correlations between electron pairs on two separate benzene molecules and leads to a minimum at 4.1 Å. Compared to the exact result, the equilibrium distance is about 0.1 Å too long and the interaction energy is underestimated by almost 37%; also the

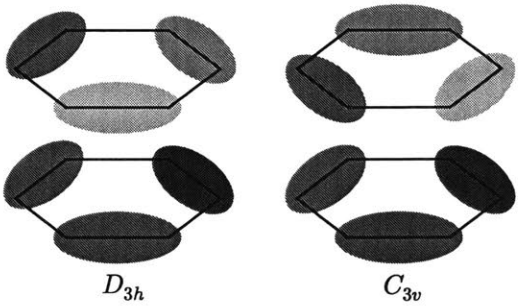


Figure 2-12: Two types of maximally localized density partitions of a benzene dimer. We assume the D_{3h} partitioning to be the default due to its higher symmetry. Numerical results indicate that the convergence of MPE strongly depends on the decomposition.

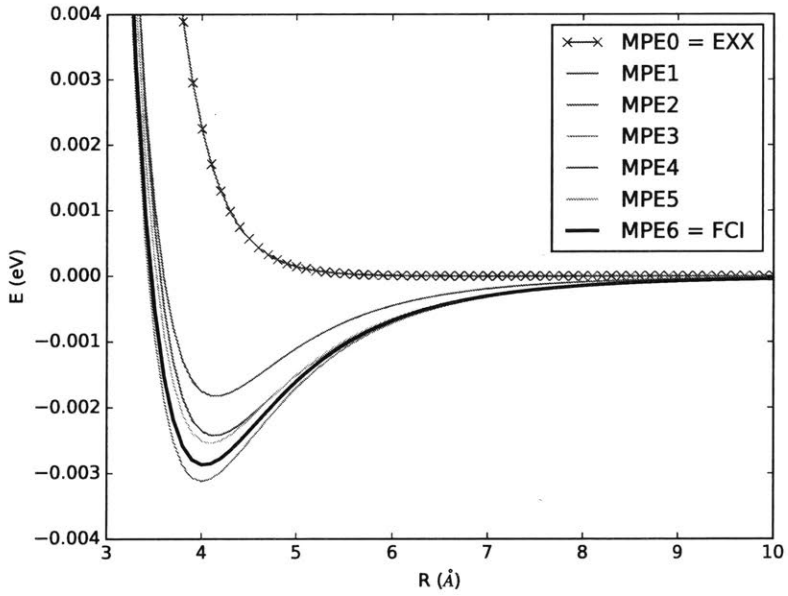


Figure 2-13: MPE0-MPE6 dissociation curves of the benzene dimer. All bonds are equivalent ($t_1 = t_2 = -2.4$ eV), density partitioning has D_{3h} symmetry.

long-range decay is too quick. MPE3 corrects this result substantially, yielding the correct equilibrium distance. The binding energy is overestimated by approximately 8% at equilibrium, but the interaction energies are very accurate beyond 6 Å. Interestingly, MPE4 gives a substantial overcorrection, shifting the minimum again to 4.1 Å and underbinding by over 15%. MPE5 is only a slight improvement, while MPE6 is a relatively significant correction leading to the exact result. Overall, already MPE2 captures the qualitative behavior of the PES and locates the equilibrium distance fairly accurately. The deterioration of MPE4

energies compared to MPE3 might be related to the fact that three pairs are conjugated within a benzene ring.

We can break the conjugation completely by setting every second transfer integral to zero (Figure 2-14a); thus generating effectively a Kekule structure of the benzene molecule. The geometry of the dimer is not changed, so all two-electron interactions in the PPP Hamiltonian are the same as in the original dimer model. In this case MPE2 leads to approximately 45% overbinding. MPE3 energies are already close to the reference, indicating that three-pair correlations can be important even in the model with broken conjugation. MPE4 and MPE5 results are visually indistinguishable from the reference values. This example confirms that in the presence of conjugated electron pairs, the convergence of MPE is likely to be more challenging.

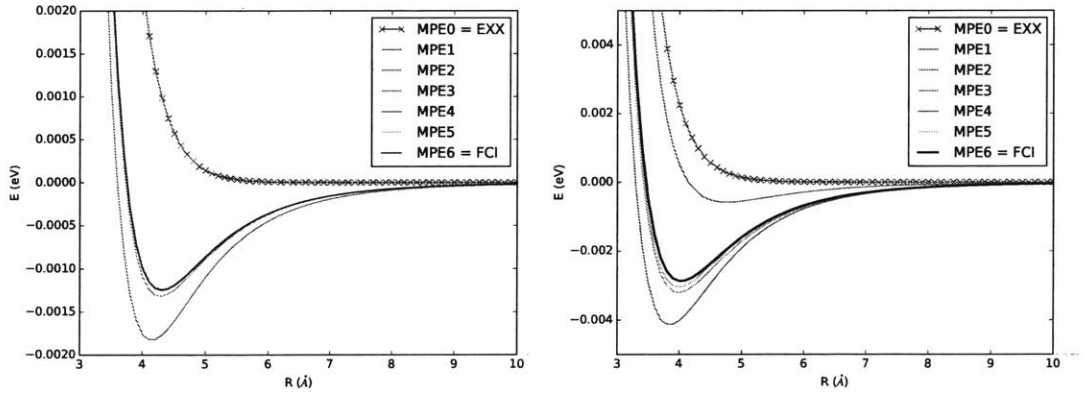


Figure 2-14: MPE0-MPE6 dissociation curves of the benzene dimer. (a) With broken conjugation within a ring. The alternating resonance integrals are $t_1 = -2.4$ eV and $t_2 = 0$ eV, density partitioning has D_{3h} symmetry. (b) With rotated density decomposition pattern. All bonds are equivalent ($t_1 = t_2 = -2.4$ eV), density partitioning has C_{3v} symmetry.

The results presented so far were obtained with pair densities on one benzene that are mirror reflection of the pair densities on the other molecule (left-hand side of Figure 2-12). However, MPE results clearly depend on the way how the total density is decomposed. To investigate this effect we consider also a pattern where pair densities on one benzene are rotated by 60° with respect to the other one (right-hand side of Figure 2-12). Comparing Figures 2-14b and 2-13 reveals that the choice of partitioning is decisive for the convergence of MPE results. For the C_{3v} partitioning (Figure 2-14b), MPE2 produces only a shallow

minimum at 4.8 Å and the binding energy is underestimated by approximately 80%. MPE3 corrects this result to a major extent, exhibiting a minimum that is only 12% too low, but MPE4, instead of getting closer to the reference, leads to even greater overbinding. MPE5 is again very close to the exact result.

The example of benzene dimer shows that the convergence of MPE interaction energies may be non-monotonic. This points to the fact that MPE may behave less well if density decomposition breaks a system of strongly delocalized electrons. In such case the results are also sensitive to the partitioning pattern. Good results obtained with MPE3 suggest that the order of the expansion needs to be somehow balanced with the number of electron pairs that are delocalized.

2.4.5 Ethylene and benzene stacks

Interaction energy is non-additive and so is the dispersion energy component [184, 185]. The non-additivity of dispersion results from dynamical screening of polarizabilities by the surrounding subsystems and from the truly many-body nature of interactions between them, which cannot be reduced to a sum of pairwise contributions. Due to its long-range character, dispersion is a collective effect, whereby the total stabilization energy grows superlinearly with the size of the system. These effects contribute to the overall cooperativity of binding of van der Waals complexes. To study how the MPE interaction energy scales with the system size, we consider stacks of ethylene and benzene molecules. To ensure equivalence of sites and thus uniform site densities, we impose cyclic boundary conditions. The distances between monomers are fixed at 4.1 Å for ethylene and 4.0 Å for benzene stacks. Figures 2-15 and 2-16 show additional stabilization per monomer due to intermolecular dispersion interactions as a function of the number of monomers in the system. To facilitate comparison, plots show relative changes with respect to the interaction energy of a trimer for every MPE level.

For the ethylene stack, MPE3 and MPE4 energies are visually indistinguishable, so MPE3 is already the converged result. Cooperativity of dispersion interactions is noticeable as additional molecules in the stack lower the stabilization energy per monomer. This monotonic decrease flats out at 7 monomers for all considered levels of MPE. This indicates that while MPE2 underestimates the additional stabilization by about 25%, it properly scales with the system size. The converged total interaction energy per monomer is 0.76 meV, so the cooperativity effect accounts for about 3% of the total value.

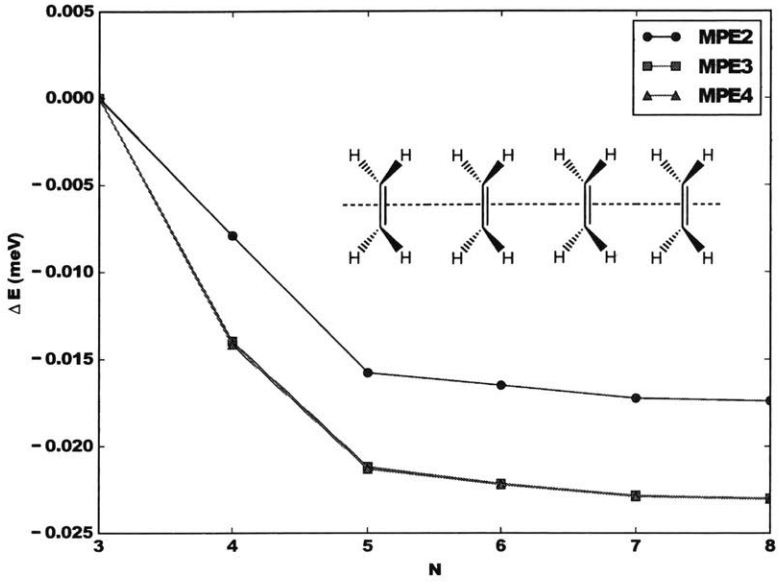


Figure 2-15: MPE2-4 stabilization energies per monomer of an ethylene stack. Inset shows the linear arrangement of ethylene molecules in the stack. The distance between monomers is 4.1 Å and cyclic boundary conditions are imposed.

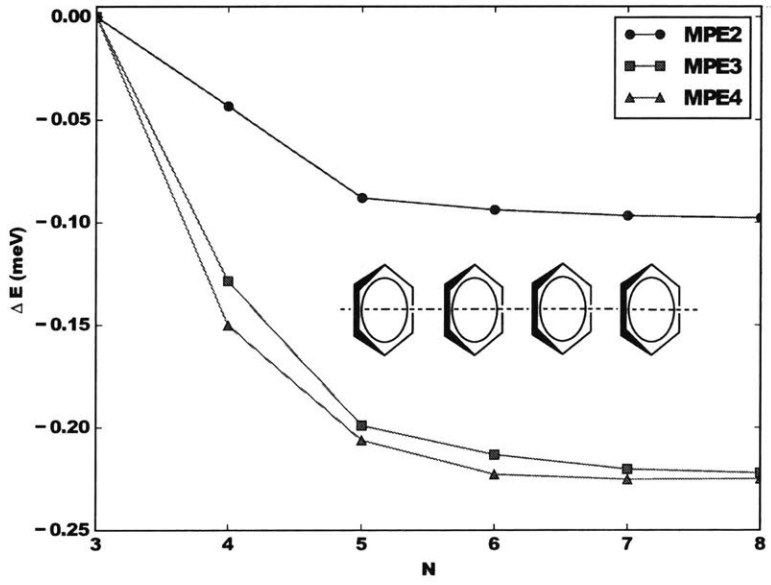


Figure 2-16: MPE2-4 stabilization energies per monomer of a benzene stack. Inset shows the linear arrangement of benzene molecules in the stack. The distance between monomers is 4.0 Å and cyclic boundary conditions are imposed.

The results for benzene stacks are qualitatively similar. The relative effect of adding MPE3 corrections is more pronounced as MPE2 recovers less than 50% of the additional stabilization. Adding MPE4 correction stabilizes the system further; however, the effect is comparatively small. For 7 and 8 monomers in the stack, some loss of accuracy becomes apparent due to accumulation of numerical errors. Contrary to the total interaction energy in the benzene dimer (Figures 2-13 and 2-14), additional stabilization due to many-body interactions between different monomers seems to converge monotonically with the level of MPE. As the total MPE4 interaction energy per monomer stabilizes at 4.7 meV, the additional stabilization due to cooperativity accounts for almost 5% of the effect.

These applications of MPE to the PPP hamiltionian clearly show that already low levels of the expansion successfully recover the dispersion interactions, which are absent in the mean field description. As all the terms are defined as density functionals, the implication is that MPE provides an efficient way to formulate a dispersion correction scheme for any class of approximate exchange-correlation functionals. The procedure is completely non-empirical and naturally accounts for both intra- and inter-molecular dispersion. As it does not assume the asymptotic $1/R^6$ dependence, it works equally well at any separation between interacting subsystems and does not require any damping at short ranges. The effect of cooperativity of many-body interactions is properly accounted for and the scheme goes beyond the pair-wise additive correction and effectively includes the dynamical dielectric screening effects. These attractive characteristics warrant pursuing further development of the proposed scheme and incorporating it in practical DFT computations.

2.5 Conclusions

In this chapter, we have developed the many-pair expansion method and shown MPE is a systematically improvable hierarchy of density functional approximations to the total energy of a quantum many-body problem. Through applications to 1D/2D Hubbard models and 1D Peierls-Hubbard model, we demonstrate MPE provides systematic improvement to the description of strongly correlated systems, which posts significant challenges for almost all approximate density functionals. Even at low levels of expansion, MPE can give accurate predictions for strongly correlated electrons in DFT. The study of more challenging Peierls-Hubbard model points out the importance of using optimal partitioning of the total density.

In the later work on PPP Hamiltonians for several π-electron systems, we reveal that energy corrections decay rapidly with the distance between pair densities. For a large system this observation allows to neglect a vast number of corrections, leading to reduced scaling at the cost of introducing small error. Additionally, we have shown that the convergence of the expansion depends on the partitioning of the total density. In particular, an optimal partitioning leads to accurate results already at low orders, making MPE calculations practical for large systems. It has been shown that already MPE2 recovers the qualitative and semi-quantitative features of dissociation curves of van der Waals complexes and recovers the $1/R^6$ long-range asymptotic behavior of dispersion energy. The importance of three-pair contributions has been shown on the example of stacked benzene molecules.

In conclusion, we have shown relatively low orders of MPE expansion are able to account for both strong and weak long-range correlations, for which common density functional approximations fail miserably. MPE thus shows promise for dealing with a wide range of applications using a universal method within the DFT framework. In this chapter we focus on model systems, while we would like to extend MPE to real molecular systems with *ab initio* Hamiltonians in the following chapters.

2.6 Acknowledgements

This work is carried out together with Dr. Piotr de Silva and Dr. Helen van Aggelen. We thank Dr. Matt Welborn for providing Bethe ansatz reference data for 1D Hubbard model and Dr. Shiwei Zhang for providing AFQMC data for 2D Hubbard model.

Chapter 3

Self-attractive Hartree decomposition

3.1 Introduction

In this chapter, we develop a new density partitioning method – self-attractive Hartree (SAH) [186], for two purposes. First, the SAH decomposition provides a unique tool to generate pair densities in the implementation of the many-pair expansion for molecular systems. The v-representability requirement for pair densities makes the implementation of molecular MPE difficult, because there is no easy way to partition the molecular density into sufficiently smooth fragments in finite Gaussian basis sets. In addition, we also require the pair densities to be localized to accelerate the convergence of MPE series so that low orders of expansion are enough for accurate results. This localization requirement is similar to employing localized molecular orbitals (LMOs) [176] in the method of increments [187]. Unfortunately, pair densities constructed out of LMOs cannot meet the v-representability criterion in MPE due to the existence of nodes. Another straightforward idea may be partitioning the density of a $2N$-electron system equally into N pair densities. In that case, the pair densities are satisfactorily smooth but highly delocalized, leading to poor convergence of MPE expansions. The proposed SAH decomposition is able to satisfy both v-representability and localization requirements to be used in molecular MPE. However, we will not further discuss the use of the SAH decomposition in MPE in this chapter but leave this topic to Chapter 4. Second, we would like to develop a tool to extract chemical bonding information directly from the electron density. This is the main focus of this chapter.

Quantum chemistry methods are widely used to explain and predict chemical phenomena by calculating the structures and properties of molecules. However, there is no direct

and unambiguous mapping between quantum chemical calculations and classical chemical concepts like chemical bonding, atomic charges, resonance and conjugation [188, 189, 190, 191, 192, 193, 194]. These chemical concepts, though mostly empirical and not observable, still play a critical role in the description and understanding of chemistry. They are closely related to a more general concept of individual electron pairs and their localization in space, which naturally cannot be rigorous in a system of many indistinguishable particles. Nevertheless, such conceptual tools are extremely useful to facilitate understanding in structures of molecules with unusual bonding [195], intermolecular interactions [196, 197], chemical reactivity [198], etc.

A natural framework to conceptualize electron pairing is the mean-field approximation to a many-electron wavefunction. Molecular orbitals (MO), which are the building blocks of a Slater determinant wavefunction, describe electrons moving in an effective field of all the others. However, canonical molecular orbitals (CMO) are typically delocalized over large fragments of the system which impairs their interpretative significance. A solution to this problem exploits the non-uniqueness of MOs in the Slater determinant and transforms CMOs to localized molecular orbitals (LMO) through a unitary transformation. Various orbital localization schemes have been proposed in the literature. They either rely on maximization of some localization measure [199, 176, 200, 201, 202] or take advantage of the localized atomic basis set to find localized and fully occupied natural orbitals (eigenvectors of the idempotent density matrix) [203, 196, 204]. In addition to conceptual advantages, LMOs have also found numerous applications in reduced scaling electronic structure methods [205, 206, 207, 208]. Nevertheless, interpretative power of orbitals has fundamental limitations. MOs are required to be mutually orthogonal which results in a complicated nodal structure which often interferes with a clear cut interpretation. Another deficiency of the orbital picture is that it does not account for electron correlation effect, which is described by interaction of different orbital configurations. While the concept of orbitals can be restored by diagonalizing the one-particle density matrix, the number of these natural orbitals with non-zero occupations far exceeds the number of electrons, limiting their interpretive value. Additionally, orbital analysis is not directly applicable to results obtained with orbital-free methods like orbital-free DFT [209] or direct density matrix optimization [210].

When looking for alternatives to the orbital analysis, it is tempting to focus on the electron density as the descriptor of a many-electron system. This is appealing since the density

is well defined in any ab initio electronic structure method and is also available experimentally from X-ray measurements. The first Hohenberg-Kohn theorem [211] guarantees that the electron density embodies all the information about the system, but extracting it is in general a daunting task. In particular disentangling different pairs of electrons and probing their localization is very difficult as contributions from different electrons are washed out. Still, there are a range of theoretical tools to analyze electron density distributions. The most prominent is perhaps the quantum theory of atoms in molecules (QTAIM) [212, 213]. In QTAIM, the physical space is divided into non-overlapping sectors separated by zero-flux density surfaces. Each sector contains one nucleus; thus it defines an atom in the molecule and its net charge. Within this theory the Laplacian of the density is also used to probe local density concentrations, which correspond to atomic shells and chemical bonds. Other scalar fields developed to analyze the electron density include localized electron locator (LED) [214], single-exponential decay detector (SEDD) [215, 216, 217] and density overlap regions indicator (DORI) [218]. An alternative way to define atoms in molecules is the Hirshfeld partitioning [73], which splits the total density based on contributions of atomic densities to the density of a promolecule. Contrary to QTAIM, the resulting atoms in molecules are overlapping and their densities are smooth functions. Another interesting approach to density partitioning is proposed by partition DFT (PDFT) [85], which introduces a local partition potential to represent the molecular density as a sum of non-interacting fragments with fractional number of electrons.

Direct decomposition of the electron density into one-electron densities, rather than atoms or molecular fragments, has not been much exploited in the literature. In this chapter, we present a new method to decompose the electron density into localized one-electron densities. Our aim is to devise an algorithm which gives subsystems that are ground states of an effective one-particle Hamiltonian. Such approach assures that fragment densities are both smooth and localized in space. We start by following the mathematical idea of maximizing the self-repulsion energy within fragment densities, similar to the Edmiston-Ruedenberg (ER) localization [199]. However, the problem is cast in a form of coupled self-consistent field (SCF) equations, by including the kinetic energy of electrons in order to regularize the solutions of the localized decomposition. The resulting non-orthogonal ground-state solutions define electron pairs and reveal chemical bonding patterns which are consistent with the non-orthogonal Valence Bond (VB) theory [219]. Unlike in the QTAIM

approach, our decomposed electron pairs are smooth and can better illustrate the chemical bonds visually. Furthermore, we apply this method to hydrogen-bonded systems to show it can not only visualize chemical bonding in molecular complexes but also accurately measure the strength of hydrogen bonding directly from the input electron density.

3.2 Theory

3.2.1 Localization of fragment densities

Given any spin-compensated molecule with $2N$ electrons, we want to decompose its electron density (which is an input density in our method) $\rho_\sigma(\mathbf{r})$ into a sum of one-electron densities $\{\rho_{i,\sigma}(\mathbf{r})\}$ such that

$$\rho_\sigma(\mathbf{r}) = \sum_i^N \rho_{i,\sigma}(\mathbf{r}), \quad \int \rho_{i,\sigma}(\mathbf{r})d\mathbf{r} = 1, \quad \rho_{i,\sigma}(\mathbf{r}) \geq 0, \quad \sigma = \{\alpha, \beta\}. \tag{3.1}$$

Since $\rho_\alpha(\mathbf{r})$ is equal to $\rho_\beta(\mathbf{r})$ in the remainder of the manuscript, we drop the spin label throughout and refer to $\rho_{\text{in}}(\mathbf{r})$ as the input density ($\rho_{\text{in}}(\mathbf{r}) = \rho_\alpha(\mathbf{r}) = \rho_\beta(\mathbf{r})$) and $\rho_i(\mathbf{r})$ as one-electron densities ($\rho_i(\mathbf{r}) = \rho_{i,\alpha}(\mathbf{r}) = \rho_{i,\beta}(\mathbf{r})$). The generalization to spin-uncompensated systems is straightforward, however the qualitative insights might not be as clear in this case.

In order to obtain localized one-electron densities, we start with the ER criterion to maximize the self-repulsion energy within the one-electron densities:

$$E_{ER}[\{\rho_i\}] = \sum_i^N \iint \frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \tag{3.2}$$

To find the maximizer of Eq. 3.2 subject to constraints in Eq. 3.1, we can express one-electron densities in terms of non-orthogonal auxiliary orbitals to assure that they are not negative

$$\rho_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2, \tag{3.3}$$

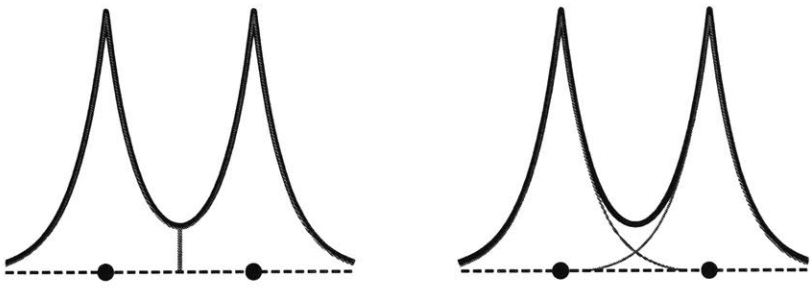
and search for a stationary point of the following Lagrangian:

$$L[\{\rho_i\}, \mu(\mathbf{r}), \{\epsilon_i\}] = \sum_i^N \iint \frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - \int \mu(\mathbf{r})\Big(\sum_i^N \rho_i(\mathbf{r}) - \rho_{\text{in}}(\mathbf{r})\Big)d\mathbf{r} + \sum_i^N \epsilon_i\Big(\int \rho_i(\mathbf{r})d\mathbf{r} - 1\Big), \tag{3.4}$$

where $\mu(\mathbf{r})$ and $\{\epsilon_i\}$ are Lagrangian multipliers. The resulting Euler-Lagrange equations have the following form:

$$2\int \frac{\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' = \mu(\mathbf{r}) - \epsilon_i. \tag{3.5}$$

However, solutions of Eq. 3.5 would not yield a satisfactory decomposition. Since we are constraining one-electron densities only to be non-negative, maximization of self-repulsion would result in fragments whose densities sharply drop to 0 outside of some regions of space. Consequently, the input density would be decomposed into a set of non-overlapping fragments. This is highly undesired from the numerical perspective, as it would make it impossible to expand the auxiliary orbitals $\{\phi_i(\mathbf{r})\}$ in a basis set of some smooth functions, e.g. Gaussians. It is also undesired from the conceptual standpoint, as we would like to think about electron pairs as localized but still overlapping and with a smooth decay of their probability densities, rather than confined to finite volumes of space. Figure 3-1 illustrates the non-smooth and smooth decompositions of the density, where the latter (Figure 3-1b) is the preferred choice. We also note that the non-overlapping fragment densities from Eq. 3.5 are different from the densities in QTAIM. QTAIM in general produces fragment densities which account for partial charge transfer and do not integrate to integers, while solutions of Eq. 3.5 always integrate to one electron.



(a) (b)

Figure 3-1: Illustrations for density partitioning. (a) Non-smooth decomposition of the density from direct maximization of self-repulsion. (b) Smooth decomposition of the density from regularized localization.

3.2.2 Regularization of the Lagrangian

A way to assure smooth decay of fragment densities is to regularize the Lagrangian in Eq. 3.4 by adding a term that penalizes any possible sharp features. This can be realized by adding the non-interacting kinetic energy as a regularization term. Consequently, we construct a regularized Lagrangian which corresponds to minimization of the sum of the kinetic energy

$$E_k[\{\phi_i\}] = \frac{1}{2}\sum_i^N \int |\nabla\phi_i(\mathbf{r})|^2 d\mathbf{r} \tag{3.6}$$

and the self-attraction energy, which is just negative of the self-repulsion (Eq. 3.2), subject to constraints in Eq. 3.1. In Eq. 3.6, E_k is the sum of kinetic energies of localized electrons rather than the real kinetic energy of the whole system. Since simultaneous minimization of the self-attraction and kinetic energy are conflicting requirements, minimization of their sum would give solutions that are regularized compared to the solutions of Eq. 3.5. We can have more control over the level of regularization by scaling the self-attraction energy by a parameter α (>0), which results in the following form of a regularized Lagrangian:

$$\begin{aligned} L[\{\phi_i\},\mu(\mathbf{r}),\{\epsilon_i\}] = \frac{1}{2}\sum_i^N \int |\nabla\phi_i(\mathbf{r})|^2 d\mathbf{r} - \alpha\sum_i^N \iint \frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ + \int \mu(\mathbf{r})\Big(\sum_i^N \rho_i(\mathbf{r}) - \rho_{\text{in}}(\mathbf{r})\Big)d\mathbf{r} - \sum_i^N \epsilon_i\Big(\int \rho_i(\mathbf{r})d\mathbf{r} - 1\Big), \end{aligned} \tag{3.7}$$

where $\rho_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2$.

Solving for stationary points of this Lagrangian $\left(\frac{\delta L}{\delta \phi_i} = 0\right)$ gives the following set of equations:

$$\Big[-\frac{1}{2}\nabla^2 - 2\alpha\int \frac{\phi_i(\mathbf{r}')^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \mu(\mathbf{r})\Big]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \tag{3.8}$$

Eq. 3.8 is a set of N equations for N auxiliary orbitals. They have a similar form to Hartree equations [42] which describe self-interacting charged particles. However, instead of a repulsive potential in the Hartree equation, we have a potential $-2\alpha\int \frac{\phi_i(\mathbf{r}')^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'$. This potential describes a rather peculiar interaction as charge distributions are now experiencing self-attraction. In addition, all the subsystems feel a common external potential $\mu(\mathbf{r})$ which constrains the sum of one-electron densities to be the input electron density. Therefore, Eq. 3.8 can be interpreted as describing self-attracting electrons in a common external field.

In analogy to the Hartree equation, Eq. 3.8 can be solved in a self-consistent procedure using standard SCF techniques. As a result, we obtain a set of localized fragment densities that sum up to the total density, therefore we name this procedure the self-attractive Hartree (SAH) decomposition of the electron density.

Meanwhile, Eq. 3.8 has another feature that each ϕ_i is a ground-state solution of a Kohn-Sham (KS) equation, meaning that each one-electron density is v-representable by design. This is similar to solving field theory equations with quartic interactions, where the symmetry is spontaneously broken, leading to many ground-state solutions [220]. This is also a very attractive feature that can be exploited in the development of DFT. For instance, such densities might mitigate some of the problems of Perdew-Zunger self-interaction correction [106], which are related to the appearance of nodes in KS orbitals [221, 222, 223, 224]. They could also be very useful in methods such as subsystem DFT [225] and wavefunction-in-DFT embedding [83, 84]. Particularly, they are crucial for the development of the many-pair expansion (MPE) [139, 145] for molecular systems.

3.2.3 Basis set implementation

To solve Eq. 3.8 using standard linear algebra techniques, we need to have a way to represent the potential $\mu(\mathbf{r})$ and to calculate its matrix elements $< i|\mu(\mathbf{r})|j >$ in some convenient basis. To this end we introduce an auxiliary basis set $\{\chi_P\}$[226, 227] composed of Gaussian-type atomic orbitals. Then the external potential $\mu(\mathbf{r})$ can be expanded using the nuclear potential and Coulomb potentials of functions in the auxiliary basis set [81]:

$$\mu(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_t b_t \widetilde{\chi}_t(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_t b_t \int \frac{\chi_t(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{3.9}$$

where $v_{nuc}(\mathbf{r})$ is the nuclear potential of the system and $\{b_t\}$ are the coefficients for the potential basis functions. In this work, we use the nuclear potential $v_{nuc}(\mathbf{r})$ in order to recover the proper density cusps at the nuclei. However, we also notice adding $v_{nuc}(\mathbf{r})$ is not necessary if the auxiliary basis set $\{\chi_P\}$ is large enough because the nuclear potential can be well represented by the potential basis set. The same basis set can be used to represent the self-attraction potential by expanding one-electron densities in the auxiliary basis set

$$\rho_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2 \approx \sum_P d_P^i \chi_P(\mathbf{r}), \tag{3.10}$$

with density fitting coefficients [228, 229]

$$d_P^i = \sum_Q (ii|Q)(Q|P)^{-1}, \tag{3.11}$$

where

$$(ii|Q) = \iint \frac{\phi_i(\mathbf{r})\phi_i(\mathbf{r})\chi_Q(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}', \tag{3.12}$$

$$(Q|P) = \iint \frac{\chi_Q(\mathbf{r})\chi_P(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \tag{3.13}$$

By inserting Eq. 3.10 and Eq. 3.9 into Eq. 3.8, we obtain the final working equation:

$$\big[-\frac{1}{2}\nabla^2 + v_{nuc}(\mathbf{r}) - 2\alpha \sum_P d_P^i \widetilde{\chi}_P(\mathbf{r}) + \sum_t b_t \widetilde{\chi}_t(\mathbf{r})\big]\phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{3.14}$$

We use a constrained SCF algorithm composed of an outer loop and an inner loop to solve Eq. 3.14 under the constraints in Eq. 3.1, see Algorithm 1. In the outer loop, we perform the normal SCF procedure for each one-electron density respectively to update $\phi_i(\mathbf{r})$ (and $\{d_P^i\}$ according to Eq. 3.11) iteratively. Note the constraint $\int \rho_i(\mathbf{r})d\mathbf{r} = 1$ is automatically fulfilled by occupying one electron in the lowest orbital of each $\phi_i(\mathbf{r})$. In the inner loop, $\{b_t\}$ is updated for all N SCF equations through the Wu-Yang potential inversion technique [77, 86] to satisfy $\sum_i^N \rho_i(\mathbf{r}) = \rho_{\text{in}}(\mathbf{r})$. This constraint is enforced by searching for $\{b_t\}$ that gives $\mathbf{W} = \mathbf{0}$, where

$$W_u = \int \widetilde{\chi}_u(\mathbf{r})\big(\sum_i^N \rho_i(\mathbf{r}) - \rho_{\text{in}}(\mathbf{r})\big)d\mathbf{r}. \tag{3.15}$$

This search can be achieved using the Newton's root-finding method, and the analytic Jacobian $\frac{\partial \mathbf{W}}{\partial \mathbf{b}}$ is available:

$$\begin{aligned}
\frac{\partial W_u}{\partial b_t} &= \sum_i^N \iint \frac{\delta W_u}{\delta \rho_i(\mathbf{r})} \frac{\delta \rho_i(\mathbf{r})}{\delta \mu(\mathbf{r}')} \frac{\partial \mu(\mathbf{r}')}{\partial b_t} d\mathbf{r}d\mathbf{r}' \\
&= \sum_i^N \iint \widetilde{\chi}_u(\mathbf{r}) \cdot 2 \sum_a^{vir} \frac{\phi_i^0(\mathbf{r})\phi_i^a(\mathbf{r})\phi_i^0(\mathbf{r}')\phi_i^a(\mathbf{r}')}{\epsilon_i^0 - \epsilon_i^a} \cdot \widetilde{\chi}_t(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\
&= 2\sum_i^N \sum_a^{vir} \frac{(\phi_i^0\phi_i^a|u)(\phi_i^0\phi_i^a|t)}{\epsilon_i^0 - \epsilon_i^a},
\end{aligned} \tag{3.16}$$

where ϕ_i^0 is the lowest and only occupied orbital for ith fragment density, ϕ_i^a is the virtual orbital for the same fragment density, ϵ_i^0 and ϵ_i^a are the occupied and virtual orbital energies.

Algorithm 1: SAH decomposition

Input : the total electron density ρ_{in}

Initialize: α: localization parameter; $\{\phi_i\}$: auxiliary orbitals for one-electron densities $\{\rho_i\}$; $\{b_t\}$: coefficients of the external potential μ; $\epsilon_1 = 10^{-6}$: SCF convergence criterion; $\epsilon_2 = 10^{-5}$: density matching criterion

while $\sum_i^N ||\rho_i^{k+1} - \rho_i^k|| > \epsilon_1$ **do**

  Compute one-electron density coefficients $\{d_P^i\}$ according to Eq. 3.11;

  while $||\frac{\partial \mathbf{W}}{\partial \mathbf{b}}|| > \epsilon_2$ **do**

    Construct N Fock matrices $\{\mathbf{F_i}\}$ according to Eq. 3.14;

    Diagonalize Fock matrices $\{\mathbf{F_i}\}$ and obtain temporary orbitals $\{\phi_i\}$;

    Compute the density matching function $\mathbf{W}$ according to Eq. 3.15;

    Update $\{b_t\}$: $\mathbf{b} = \mathbf{b} + \left(\frac{\partial \mathbf{W}}{\partial \mathbf{b}}\right)^{-1} \cdot \mathbf{W}$, acoording to Eqs. 3.15-3.16;

  end

  Construct updated Fock matrices $\{\mathbf{F_i^{k+1}}\}$;

  Obtain new orbitals $\{\phi_i^{k+1}\}$ and one-electron densities $\{\rho_i^{k+1}\}$;

end

3.3 Computational details

For simplicity, we only use the electron density from quantum chemical calculations as the input. All quantum chemical calculations are done using Q-Chem 4.2 software package [230]. The ground state geometries are first optimized using the second-order Møller-Plesset (MP2) method [34] for all single molecules in cc-pVTZ basis set [231] and hydrogen-bonded dimers in aug-cc-pVTZ basis set [232]. Unless otherwise specified, we then obtain the electron density from DFT calculations with B3LYP functional [49] for single molecules in uncontracted cc-pVTZ (u-cc-pVTZ) basis set and hydrogen-bonded dimers in uncontracted aug-cc-pVTZ (u-aug-cc-pVTZ) basis set. We do not use the electron density from correlated wavefunction methods in this paper so the electron correlation effects are not covered. However, the formalism of SAH decomposition allows for analysis of correlated electron densities and this will be investigated in the future work.

For the SAH decomposition, u-cc-pVTZ and u-aug-cc-pVTZ basis sets are used for single molecules and hydrogen-bonded dimers respectively to expand orbitals $\{\phi_i(\mathbf{r})\}$ in Eq. 3.14. Meanwhile, the auxiliary basis set $\{\chi_P(\mathbf{r})\}$ for expanding potentials in Eq. 3.14 needs to be constructed carefully. As known from the literature, the potential inversion and optimized

effective potential (OEP) techniques are numerically unstable in finite basis sets [233, 234]. Several solutions exist to avoid ill-conditioning problems in potential inversions [235, 236, 237, 82]. Here, we follow the work of Görling and co-workers [235], to balance the orbital and auxiliary basis sets. The auxiliary basis set is constructed out of the orbital basis by removing some of the most compact and diffuse functions. The detailed description of balanced orbital and auxiliary basis sets can be found in the Appendix. VMD [238] and Molden [239] are used to generate fragment density pictures.

3.4 Results: chemical bonding from SAH

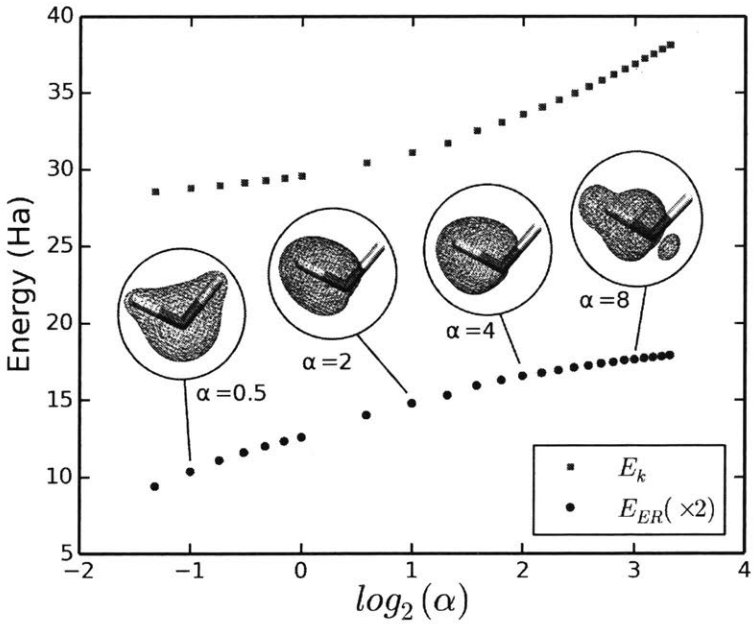


Figure 3-2: Analysis of the self-repulsion energy (E_{ER}) and the kinetic energy (E_k) of fragment densities in a water molecule. The least localized one-electron density (isovalue=0.025 a.u.) is shown for different α values.

We first perform a numerical analysis of SAH decomposition for the water molecule in order to make the best choice of the parameter α in Eq. 3.14. Figure 3-2 shows the self-repulsion energy (E_{ER}, see Eq. 3.2) and the kinetic energy (E_k, see Eq. 3.6) of all fragment densities in H_2O as a function of the parameter α. As mentioned before, α controls the ratio of E_{ER} to E_k in the Lagrangian, and thus plays an important role in balancing the extent of localization and smoothness. As can be seen, when α increases, the optimized

one-electron densities indeed have larger E_{ER} and E_k, indicating the one-electron densities are more localized but less smooth.

To show how fragment densities look like as α changes, we also plot the least localized one-electron density in H_2O for different α values in Figure 3-2. Starting from small α values (e.g., $\alpha = 0.5$), the E_k term is dominating in the optimization of the Lagrangian. Hence, the one-electron densities are very smooth but delocalized, and the least localized one still resembles the total density of a water molecule. It is obvious that $\alpha = 0.5$ is too small to give localized and meaningful fragment densities. When α is increased to 2, the E_{ER} and E_k terms are both playing important roles in the process of the Lagrangian optimization. One can see in Figure 3-2 that the least localized one-electron density at $\alpha = 2$ is both localized and smooth. Particularly, it actually resembles the O-H bond, suggesting we can already extract useful chemical bonding information by using $\alpha = 2$. As α is further increased to 4, the least localized one-electron density does not change much: it is a little more localized and still looks smooth. However, it is not the case when we have a even higher value $\alpha = 8$. When α is too large, the role of the regularization term E_k fades away, which results in the irregular shapes and nodes in fragment densities. Note that the existence of nodes is entirely due to the use of a finite basis for both orbitals and potentials. In a complete basis set limit, no nodes should exist because all one-electron densities are ground-state solutions by construction. It has been found that the presence of spurious nodes in the lowest-lying orbitals is rather a rule even in regular Kohn-Sham calculations with atomic basis sets [240]. From this analysis, using α values from 2 to 4 is reasonable for generating both localized and smooth one-electron densities in H_2O.

We further present a more detailed analysis of our partitioning approach for six molecules in Figure 3-3, in which all normal bonding patterns are included. Instead of E_{ER} and E_k, we plot E_{ER}/E_k as a function of $log_2(\alpha)$. Since we want to have a large E_{ER} and a small E_k at the same time, computing E_{ER}/E_k may be helpful to pick the best α, with which E_{ER}/E_k should have the maximum value. As shown in Figure 3-3, the E_{ER}/E_k curves for CH_4, N_2, C_2H_4 and H_2O have maxima around $\alpha = 3 \sim 4$. However, for molecules that have heavier atoms, larger α values are needed to achieve the maximum E_{ER}/E_k. For example, from H_2O to H_2S to H_2Se, the α value for maximum E_{ER}/E_k increases from $\alpha = 4$ to $\alpha = 7$ to $\alpha \geq 10$.

To choose one unified α parameter for all molecules, we take a more detailed look at the

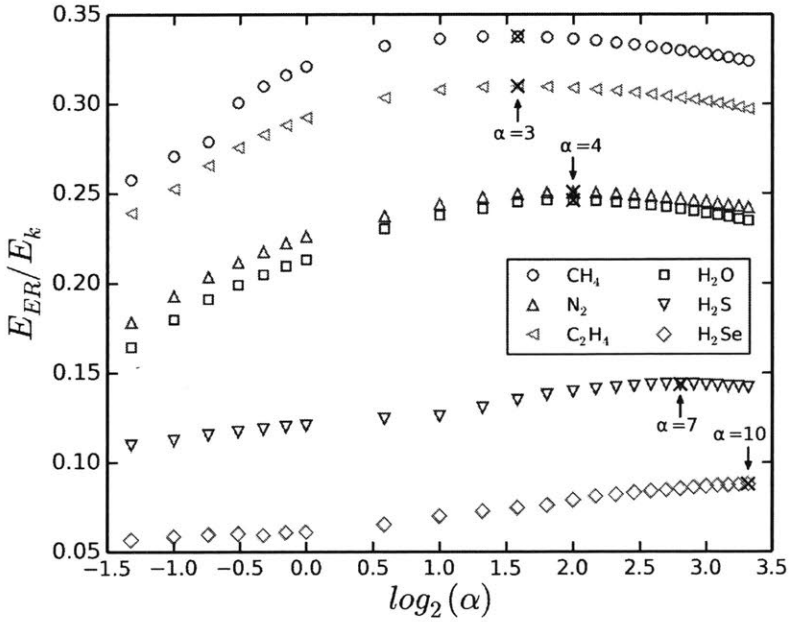


Figure 3-3: E_{ER}/E_k for six molecules. The maximum E_{ER}/E_k for each molecule is shown by a black cross mark.

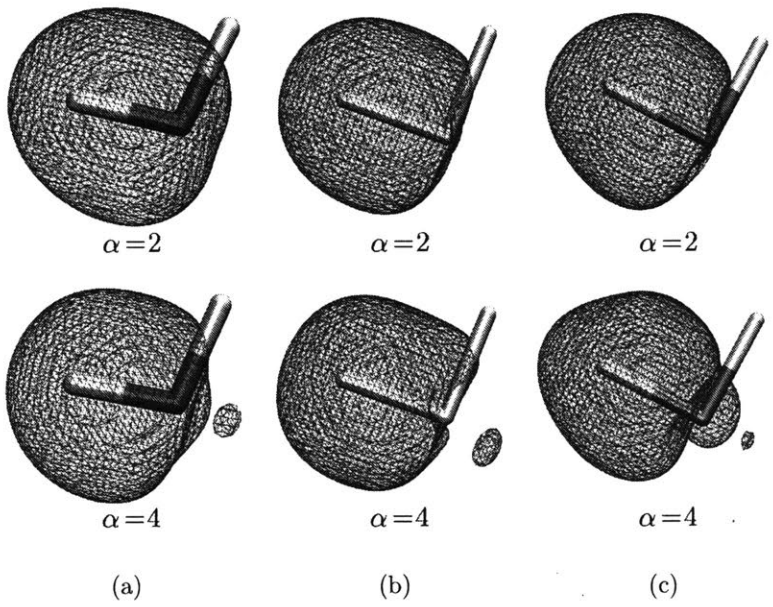
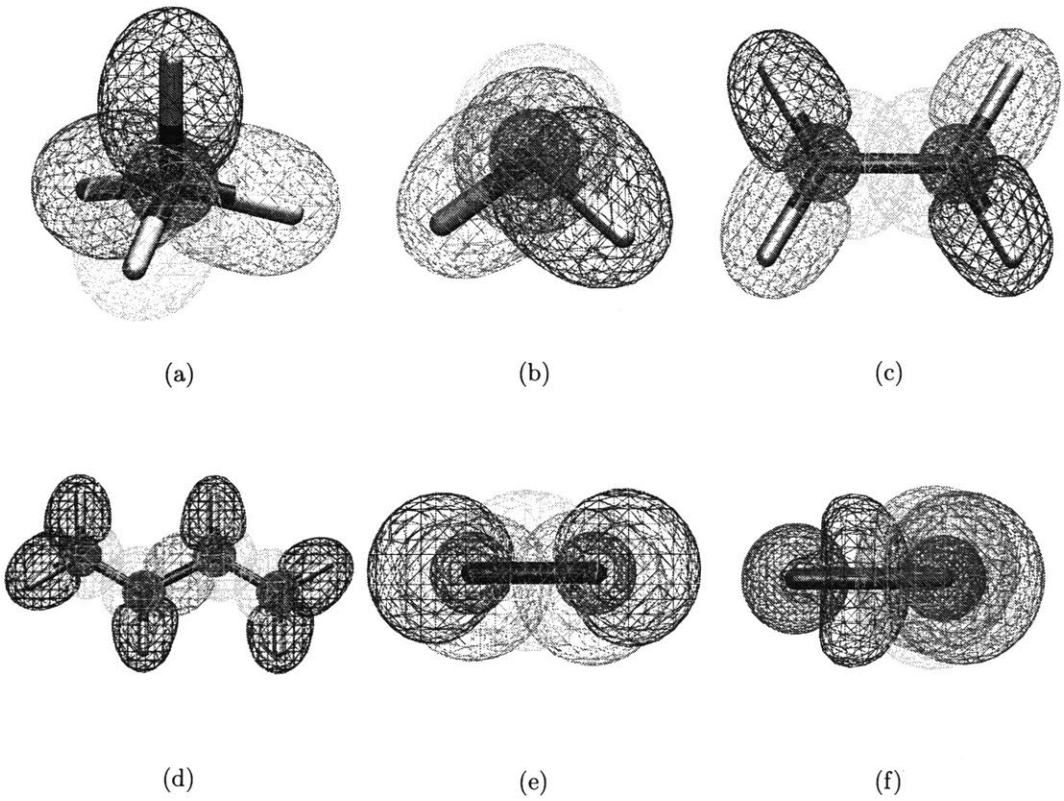


Figure 3-4: Comparison of the least localized one-electron densities generated using different α values for three molecules: (a) H_2O (isovalue = 8×10^{-3} a.u.); (b) H_2S (isovalue = 4×10^{-3} a.u.); (c) H_2Se (isovalue = 4×10^{-3} a.u.).

one-electron densities for H_2O, H_2S and H_2Se. As shown in Figure 3-4, the least localized one-electron densities generated using $\alpha = 2$ are localized and smooth for all three tested molecules. However, when using $\alpha = 4$, the one-electron densities have nodes and are not perfectly smooth, no matter what maximum α values these molecules have in Figure 3-3. This pattern has been observed for all tested molecules and 2 is the largest value for α that does not produce obvious nodes. Therefore, we choose $\alpha = 2$ for all remaining molecular systems analyzed in this article. We also note that for much heavier atoms than the tested ones the optimum α values may change and need to be chosen carefully.



(a) (b) (c)

(d) (e) (f)

Figure 3-5: Localized and smooth pair densities of six molecules ($\alpha = 2$, isovalue = 0.03 a.u. for a-e, isovalue = 0.02 a.u. for f): (a) CH_4, (b) H_2O, (c) C_2H_4, (d) 1,3-Butadiene, (e) N_2, (f) BeO.

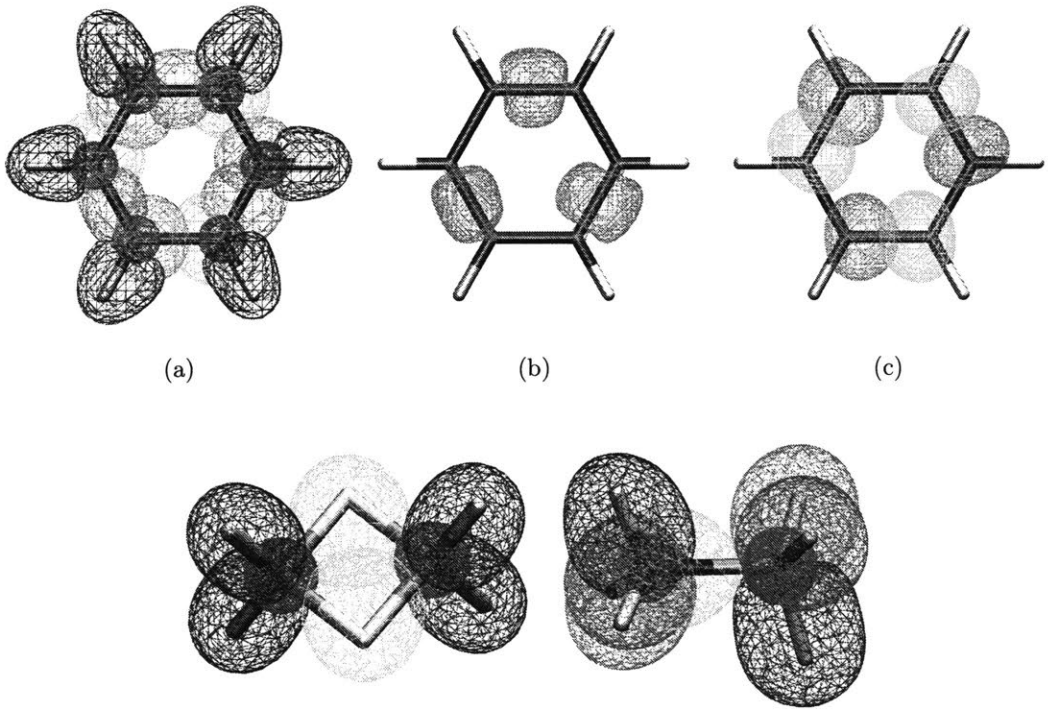
Since the spin-up one-electron densities are the same as their spin-down counterparts, they naturally form localized electron pairs. Such pair densities are very useful in illustrating chemical bonds, as shown in Figure 3-5. The first molecule tested is CH_4 (Figure 3-5a). As can be seen, the total electron density of CH_4 is decomposed into 5 localized and smooth pair densities. One of them is a core pair density on the carbon atom and the others are four

C-H single-bonding pair densities. This bonding pattern is in agreement with VB theory: the covalent bond is formed between two atoms where each atom contributes one electron. Similarly, it is shown in Figure 3-5b that the total density of H_2O is partitioned into 5 pair densities, including one core pair density and two O-H single-bonding pair densities. Moreover, there are two more pair densities located around the oxygen atom, which are recognized as the non-bonding lone pairs. As shown in this example, one strength of SAH decomposition is it explicitly breaks symmetry so that lone pairs in agreement with chemical intuition can be extracted directly from the electron density. This result is similar to Boys and ER localizations, but different from Pipek-Mezey localization or NBO analysis, in which the σ and π symmetry is preserved for lone pairs [241, 242, 243]. Another strength of our method is that, unlike other density-based methods such as QTAIM, it is more useful in visualizing chemical bonds by generating very smooth pair densities.

The bonding situation becomes more complicated when it comes to C_2H_4 (Figure 3-5c). Among the 8 resulting pair densities, two core pair densities and four C-H single bonds are apparently identified. The other two overlapping pair densities (green) sitting between two carbon atoms constitute the C=C double bond. This is reminiscent of the VB picture of a "banana bond" that results from two equivalent tetrahedral orbitals from each atom, originally proposed by Pauling [244]. We see it as a natural result of our design principle because decomposing the double bond into two pair densities from the middle guarantees the pair densities are nodeless and localized simultaneously. We further apply our partitioning scheme to 1,3-butadiene to test its performance for the linear conjugated molecules. As shown in Figure 3-5d, the bonding pattern in 1,3-butadiene is clearly revealed: two C=C double bonds (green) on 1,3-positions and one C-C single bond (orange) between the middle two carbon atoms. In Figure 3-5e, the triple bond pattern is shown for N_2. This triple bond is formed by a σ-bonding pair density (green) and two pair densities (orange) similar to the double bonding densities in Figures 3-5c and 3-5d. Two blue pair densities represent the lone pairs for two nitrogen atoms.

In addition to covalent bonding, SAH decomposition is also able to reveal ionic bonding. Figure 3-5f is an example. In BeO, the beryllium atom loses its electrons to the oxygen atom and these atoms form the ionic bond through the electrostatic attraction. Decomposing the electron density of BeO, there is one pair density (blue) that shows strong charge transfer character. It completely comes from the beryllium atom but is now shared between the two

atoms, which is a sign of ionic bonding. However, this pair density is not fully located on the oxygen atom, suggesting that these two electrons are not completely transferred from the beryllium atom to the oxygen atom. This result agrees well with the partial charge study using the Hirshfeld analysis [245].



(a) (b) (c)

(d) (e)

Figure 3-6: Localized and smooth pair densities of three molecules with nontrivial bonding patterns ($\alpha = 2$). (a) Benzene (isovalue = 0.03 a.u.). (b) Single bonding pair densities of benzene (isovalue = 0.03 a.u.). (c) Double bonding pair densities of benzene (isovalue = 0.03 a.u.). Two pair densities contribute to one double bond. (d) B_2H_6 (isovalue = 0.025 a.u.). (e) BH_3NH_3 (isovalue = 0.015 a.u.).

The more interesting cases are to apply SAH decomposition to some nontrivial bonding situations. Figure 3-6a shows the bonding patterns in the benzene molecule. The pair densities look like a cyclohexatriene structure, also known as the Kekulé structure, where the C-C single bonds (orange, Figure 3-6b) and C=C double bonds (one green and one cyan pair densities constitute one double bond, Figure 3-6c) alternate in the benzene ring. This again agrees with VB theory, which explains the benzene structure as a resonance between two Kekulé structures. Our individual pair densities break the symmetry in benzene, while

they still add up exactly to the total symmetric electron density. This feature is different from LMO methods, in which the symmetry is normally maintained. It is also different from typical VB theory, in which single resonance structure does not have symmetric total density. Comparing Figure 3-6a with Figure 3-5d, one realizes that the C=C double-bonding pair densities have a larger overlap with the C-C single-bonding pair densities in the benzene molecule, making sure the electron densities between every two carbon atoms are same.

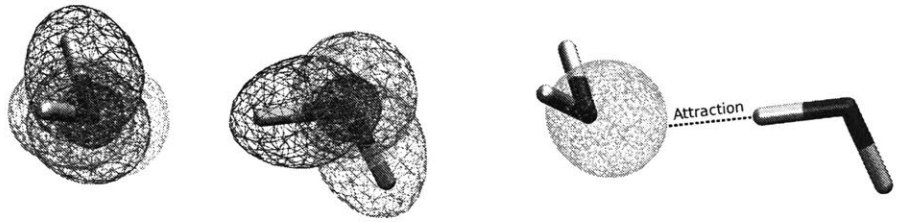
The other nontrivial bonding molecules we study are B_2H_6 and BH_3NH_3. These two molecules have unusual bonding due to the electron-deficient nature of boron. The bonding between the two boron atoms in B_2H_6 is mediated through a bridging hydrogen atom, which leads to a 3-center 2-electron bond. As shown in Figure 3-6b, the two 3-center 2-electron bonds (green) in B_2H_6 are successfully revealed. In BH_3NH_3, borane can bind with ammonia through a dative bond (or a charge-transfer bond). We show this bonding pattern in Figure 3-6c. The green pair density completely comes from the nitrogen atom (as a lone pair) but now serves as an N-B single bond.

3.5 Application: hydrogen bonding

We have demonstrated that SAH decomposition can clearly reveal chemical bonding in single molecules. Here, we further apply it to hydrogen-bonded dimer systems, to show that this method can be a useful tool to analyze non-covalent interactions as well. Hydrogen bonding is one of the most significant non-covalent interactions, and being able to measure the strength of a hydrogen bond is crucial for understanding the related chemical properties, such as the structures of DNAs and proteins. Many quantum chemical calculations have been performed to explain the hydrogen bonding strength by computing the binding energy of the hydrogen-bonded systems [246, 247, 248]. However, direct calculations are not applicable when multiple or intramolecular hydrogen bonds exist and special treatment of the system is needed [249, 250, 251, 252]. Therefore, our goal is to develop a method for the local measurement of the hydrogen bonding strength based on our SAH pair densities.

3.5.1 Hydrogen bonding strength indicator

We first use the water dimer as an example in Figure 3-7a to show the density decomposition. The total electron density of a water dimer is decomposed into 10 pair densities, to which



(a) (b)

Figure 3-7: (a) SAH decomposition ($\alpha = 2$, isovalue = 0.03 a.u.) of a water dimer. (b) The hydrogen bond is defined as electrostatic attraction between a lone pair density in the proton-accepting water monomer and a hydrogen nucleus in the proton-donating water monomer.

each water monomer contributes 5 pair densities. From our method, the concept of a water monomer can be easily recovered from the electron density of a water cluster. More importantly, the hydrogen bond in the water dimer can be identified. According to the IUPAC definition [253], the hydrogen bond is formed between a hydrogen atom (that is covalently bound to a highly electronegative atom) and the lone pair of another highly electronegative atom. Hence, we define the hydrogen bond as the interaction between a hydrogen atom in the proton-donating molecule and a lone pair density in the proton-accepting molecule, shown in Figure 3-7b. It is then interesting to investigate whether the classical electrostatic attraction between the lone pair density and the hydrogen nucleus can be employed as a hydrogen bonding strength indicator.

To confirm our hypothesis, we apply the density partitioning scheme to 13 hydrogen-bonded dimers (proton donor$\cdots$proton acceptor): $H_2O\cdots H_2O$, $H_2O\cdots NH_3$, $NH_3\cdots NH_3$, $HF\cdots HF$, $HF\cdots NH_3$, $HF\cdots H_2O$, $HF\cdots HCN$, $CH_4\cdots NH_3$, $NH_4^+\cdots H_2O$, $H_3O^+\cdots H_2O$, $H_2O\cdots OH^-$, $H_2O\cdots CN^-$, $H_2O\cdots CCH^-$. We compute the electrostatic attraction energy to measure the hydrogen bonding strength:

$$E_{attr} = \left| \int \rho_{LP}(\mathbf{r}) V_H(\mathbf{r}) d\mathbf{r} \right|, \tag{3.17}$$

where $\rho_{LP}(\mathbf{r})$ is the lone pair density on the proton acceptor and $V_H(\mathbf{r})$ is the nuclear potential of the hydrogen atom participating in the hydrogen bonding. When there is more than one lone pair density on the proton acceptor like H_2O and HF, we choose the lone pair

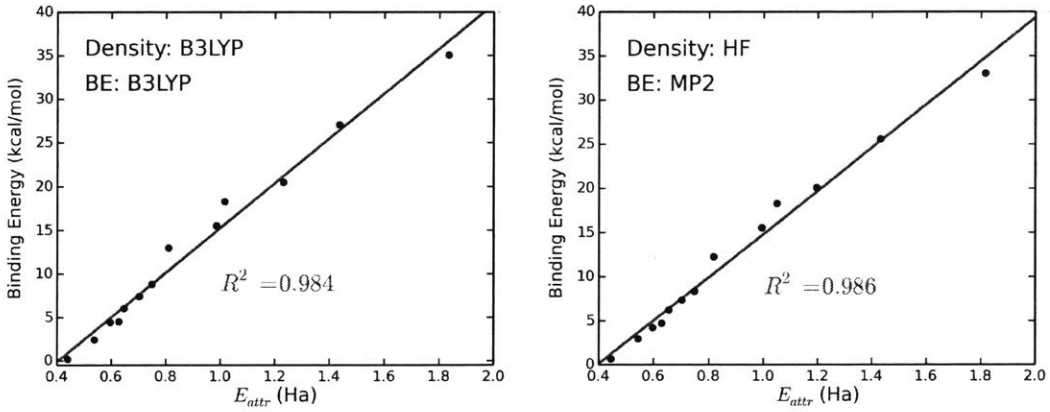
that gives largest E_{attr}. For charged hydrogen-bonded systems, we put the excess charge on the central atom and take it into consideration when computing E_{attr}. For example, the positive charge in $NH_4^+ \cdots H_2O$ is placed on the nitrogen atom, while the negative charge in $H_2O \cdots CN^-$ is placed on the carbon atom. Then for positively charged systems we add the attraction between the excess positive charge and the lone pair to Eq. 3.17:

$$E_{attr} = \left| \int \rho_{LP}(\mathbf{r}) V_H(\mathbf{r}) d\mathbf{r} + \int \rho_{LP}(\mathbf{r}) V_+(\mathbf{r}) d\mathbf{r} \right|, \tag{3.18}$$

where $V_+(\mathbf{r})$ is the potential of the positive charge. Similarly, for negatively charged systems we add the attraction between the excess negative charge and the hydrogen nucleus to Eq. 3.17:

$$E_{attr} = \left| \int \rho_{LP}(\mathbf{r}) V_H(\mathbf{r}) d\mathbf{r} + \int \rho_-(\mathbf{r}) V_H(\mathbf{r}) d\mathbf{r} \right|, \tag{3.19}$$

where $\rho_-(\mathbf{r})$ is the electron density of the negative charge.



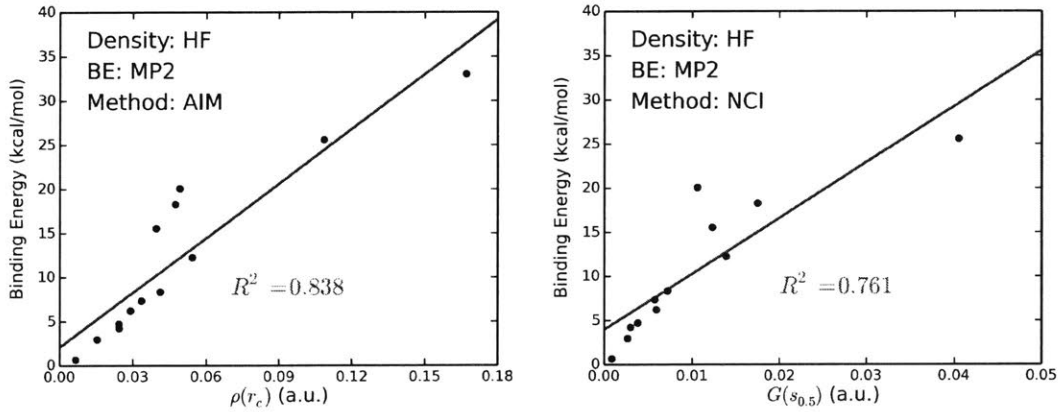
(a) (b)

Figure 3-8: The correlation between the binding energy and the electrostatic attraction energy defined in Eqs. 3.17-3.19 for 13 hydrogen-bonded dimers. (a) The B3LYP binding energy correlates with the attraction energy from the B3LYP density. (b) The MP2 binding energy correlates with the attraction energy from the HF density.

To examine whether this simple prescription can indicate the strength of hydrogen bonding, we compare E_{attr} (Eqs. 3.17-3.19) with the binding energy (BE) of the dimer systems in Figure 3-8. The BE is calculated using DFT with B3LYP functional in Figure 3-8a and MP2 in Figure 3-8b, both with the counterpoise correction [254] to account for the basis set superposition error. E_{attr} in Figure 3-8a is computed using B3LYP calculated electron

density as the input for SAH, while in Figure 3-8b the input is HF electron density. As can be seen in Figure 3-8a, there is a very strong correlation ($R^2 = 0.984$) between E_{attr} and the BE, suggesting the electrostatic attraction between the SAH lone pair and the hydrogen nucleus is a good indicator of the BE. It might be more surprising to find that E_{attr} calculated from decomposed Hartree-Fock densities correlates with MP2 binding energy equally well ($R^2 = 0.986$) as shown in Figure 3-8b. This correlation means the very accurate hydrogen bonding strength from MP2 is recovered from the electron density computed by a low level of theory (HF). Our results indicate that the electrostatic interaction is indeed dominant in hydrogen bonding [255, 256] and SAH pair densities are very useful in capturing two main related effects: (1) the electrostatic interaction between the lone pair and the proton; (2) the sharing of the lone pair electrons between the electronegative atom and the hydrogen atom. In summary, through the use of our SAH decomposition, one can identify hydrogen bonds in a complicated molecular system and estimate the strength of these hydrogen bonds directly from the electron density.

3.5.2 Comparison with other methods



(a) (b)

Figure 3-9: (a) The correlation between the MP2 binding energy and the electron density at the hydrogen bond critical point $\rho(r_c)$. (b) The correlation between the MP2 binding energy and the kinetic energy density integrated within the reduced density gradient volume $G(s_{0.5})$.

We further compare our SAH-based hydrogen bonding strength indicator (HBSI) with two other methods. The first method is from Parthasarathi et al. [251] and based on Bader's

QTAIM. The electron density at the hydrogen bond critical point ($\rho(r_c)$) was shown to correlate well with the binding energy so it can be used as a HBSI (denoted by AIM-HBSI). The other method[252] we consider here is based on the non-covalent interactions (NCI) index [257]. Lane et al. [252] found that the hydrogen-bonded OH-stretching red shifts correlate strongly with the kinetic energy density integrated within the reduced density gradient volume ($G(s_{0.5})$) that describes a hydrogen bond. This method is denoted by NCI-HBSI. These two methods are both density-based so it is interesting to compare them with SAH-HBSI. We computed $\rho(r_c)$ and $G(s_{0.5})$ for 13 hydrogen-bonded dimers tested above using Multiwfn [258] and NCImilano [259]. The HF/u-aug-cc-pVTZ density is used as the input and the results are shown in Figure 3-9.

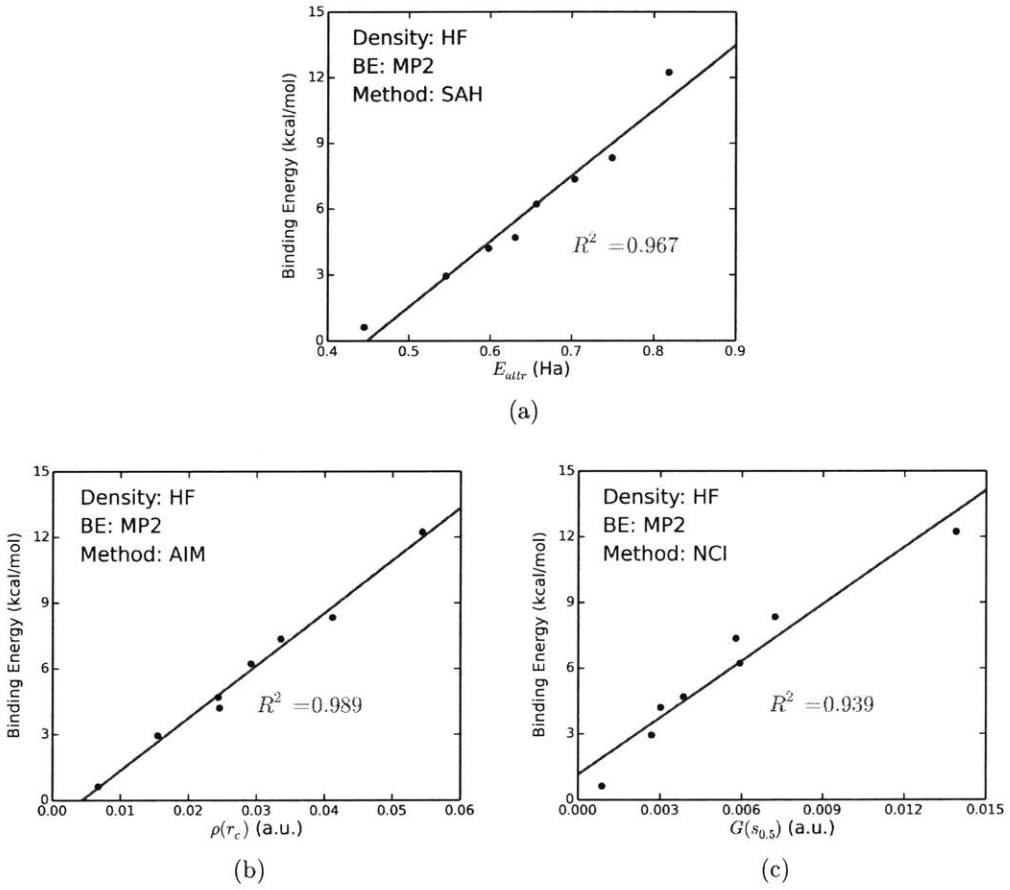
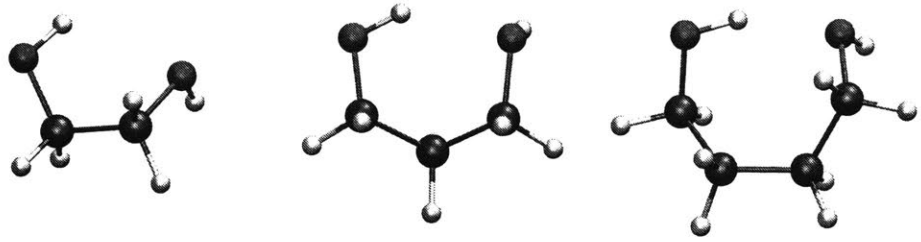


Figure 3-10: The performance of three HBSIs on 8 neutral hydrogen-bonded dimers. (a) SAH-HBSI. (b) AIM-HBSI. (c) NCI-HBSI.

Comparing Figure 3-8b with Figure 3-9, it can be seen that our SAH-HBSI has an

obviously better correlation ($R^2 = 0.986$) with the binding energy than AIM-HBSI ($R^2 = 0.838$) and NCI-HBSI ($R^2 = 0.761$). The main difference is for the charged hydrogen-bonded dimers. Since we explicitly include the electrostatic interaction due to the excess charge as shown in Eqs. 3.18-3.19, the binding energies of charged hydrogen-bonded dimers are correctly described. In AIM-HBSI and NCI-HBSI, only the electron density information in the hydrogen bond region is utilized and no such correction exists. Thus, these two methods fail to predict the hydrogen bonding strength for charged hydrogen-bonded dimers. However, we note that all three methods perform similarly well for neutral hydrogen-bonded dimers, as presented in Figure 3-10.

3.5.3 Intramolecular hydrogen bonding



(a) (b) (c)

Figure 3-11: The structures of 1,n-alkanediols: (a) 1,2-ethanediol (ED); (b) 1,3-propanediol (PD); (c) 1,4-butanediol (BD).

To demonstrate the usefulness of our approach on more complicated systems, we apply SAH-HBSI to a series of 1,n-alkanediols: 1,2-ethanediol (ED), 1,3-propanediol (PD) and 1,4-butanediol (BD) (see Figure 3-11). Intramolecular hydrogen bonds exist in these molecules and direct binding energy calculations are not applicable for estimating the hydrogen bonding strength. Previous experimental and theoretical studies [260, 261, 262] show that the intramolecular bonding strength increases as the alkane chain length increases.

We perform SAH decomposition on 1,n-alkanediol series using the electron density computed at the HF/u-aug-cc-pVTZ level and the geometries optimized at the MP2/aug-cc-pVTZ level. We compute the electrostatic attraction energy (E_{attr}) between the lone pair density and the hydrogen nucleus involved in the intramolecular hydrogen bonds based on Eq. 3.17. Using the calculated E_{attr}, we can estimate the MP2 binding energy from the

Molecule	E_{attr} (Ha)	SAH-estimated Binding Energy (kcal/mol)	$\Delta\nu$ (cm^{-1})[260]
ED	0.494	1.66	47
PD	0.581	3.92	71
BD	0.657	5.91	157

Table 3.1: Hydrogen bonding strengths of 1,n-alkanediols series based on SAH-HBSI. E_{attr} is computed according to Eq. 3.17 using HF/u-aug-cc-pVTZ densities and MP2/aug-cc-pVTZ geometries. The MP2 binding energies are estimated from the linear fit for neutral H-bonded dimers in Figure 3-8b. $\Delta\nu$ is the red shift of the first OH-stretching transition of 1,n-alkanediols compared to free OH groups, calculated using the CCSD(T)-F12a/cc-pVDZ-F12 method.

linear fit for neutral hydrogen-bonded dimers in Figure 3-8b. The results are presented in Table 3.1. The red shifts ($\Delta\nu$) of the first OH-stretching transition of 1,n-alkanediols compared to free OH groups in Ref. [260] are also presented. One can see SAH-HBSI sucessfully predicts the increasing strength of the intramolecular hydrogen bonding as the alkane chain length increases. More interestingly, the SAH-estimated MP2 binding energies can be used to directly compare the strength between different hydrogen bonds. For example, the estimated MP2 binding energy for BD (5.91 kcal/mol) is larger than the binding energy of the water dimer (4.69 kcal/mol), indicating BD's stronger hydrogen bonding strength.

3.6 Conclusions

In this chapter, we present a new decomposition method for deciphering chemical bonding from the electron density. The proposed SAH decomposition adopts the idea of minimizing self-attraction energies and regularizing the shapes of one-electron densities simultaneously, and can be solved by an efficient constrained-SCF algorithm. We show that our approach successfully illustrates the chemical bonding patterns in single molecules and molecular complexes, which agree well with non-orthogonal VB theory. We further apply it to hydrogen-bonded systems and propose a simple way to accurately measure the hydrogen bonding strength. The SAH decomposition is purely density-based, making it easy to be combined with any quantum chemistry methods. Another feature is it generates smooth and nodeless pair densities by construction. Thus it is more suitable for visualizing the chemical bonding than other density-based methods.

The future applications should focus on three aspects. First, the nature of our method

makes it effective in revealing fragment properties in complex systems. For example, we would like to study properties of water monomers in liquid water, such as their charge, dipole and polarizability [263]. Besides, we would also like to investigate the hydrogen bonding formation and strength in biological systems, such as proteins and DNAs, based on the proposed hydrogen bonding strength indicator. Second, as SAH decomposition can be used with any input electron densities, it would be attractive to combine it with X-ray measurements to extract chemical bonding information directly from experimental densities. Third, due to the smooth (v-representable) nature of the resulting fragment densities, this method is a good starting point for the many-pair expansion method [139, 145], which we will further discuss in Chapter 4. This important feature of v-representability also makes it interesting to be adopted in other DFT methods, such as self-interaction correction [106] and subsystem DFT [225].

3.7 Acknowledgements

This work is carried out together with Dr. Piotr de Silva.

Chapter 4

Many-pair expansion: towards real molecules

4.1 Introduction

In Chapter 2, we have proposed the many-pair expansion (MPE) [139, 145], which is a hierarchy of density functional approximations that systematically corrects any deficiencies of an approximate functional to finally converge to the true ground-state energy. We have shown that MPE gives accurate results for several important lattice models, including 1D/2D Hubbard [140] and 1D Peierls-Hubbard [141] models, indicating its capability of correcting strong correlation and self-interaction errors. In addition, we have applied MPE to polyacetylene as well as stacks of ethylene and benzene molecules in the Pariser-Parr-Pople (PPP) lattice model [142, 144] and demonstrated it deals very well with long-range Coulomb interactions [145]. In Chapter 2, the usefulness of MPE has only been proved in model Hamiltonians. Therefore, in this chapter, we extend MPE to real molecules by implementing it using *ab initio* Hamiltonians in Gaussian basis sets. By employing the self-attractive Hartree (SAH) decomposition [186] developed in Chapter 3, we obtain localized and v-representable density fragments, which allows us to perform numerically stable MPE calculations for molecules. We test the performance of MPE for the thermochemistry of a few small molecules and hydrogen chains as well as the symmetric bond breaking of hydrogen rings, and show that low orders of expansion are sufficient to obtain accurate results. Furthermore, we show that this method is numerically convergent for various systems

and problems tested, suggesting its potential to serve as a systemically improvable density functional hierarchy for molecular systems.

4.2 Theory

4.2.1 Implementation of MPE for molecules

The general formalisms of the many-pair expansion and self-attractive Hartree decomposition have already been discussed in Chapter 2 and Chapter 3. Here, we continue to discuss how to use SAH-generated pair densities in MPE calculations for molecules. From the SAH decomposition, we construct relevant fragment densities $\rho_q(\mathbf{r})$. Then we compute the DFT energy $E_a[\rho_q]$ as described in Eqs. 2.3-2.8, using Wu-Yang potential inversion algorithm. Similar to Eq. 3.9, we expand the Lagrange multiplier $v_s(\mathbf{r})$ in Eq. 2.4 using Coulomb potentials of functions in the auxiliary basis set $\{\chi_P(\mathbf{r})\}$:

$$v_s(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_P c_P \widetilde{\chi}_P(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_P c_P \int \frac{\chi_P(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \tag{4.1}$$

where $\{c_P\}$ are the coefficients for the potential basis functions. Then Eq. 2.4 becomes

$$L_{KS}[\{\phi_k\},\{c_P\}] = \langle\Phi| - \frac{1}{2}\nabla^2|\Phi\rangle + \int \Big(v_{nuc}(\mathbf{r}) + \sum_P c_P \widetilde{\chi}_P(\mathbf{r})\Big)\big(\rho[v_s](\mathbf{r}) - \rho_q(\mathbf{r})\big)d\mathbf{r}, \tag{4.2}$$

and each KS orbital $\phi_k(\mathbf{r})$ satisfies the equation

$$\Big(-\frac{1}{2}\nabla^2 + v_{nuc}(\mathbf{r}) + \sum_P c_P \widetilde{\chi}_P(\mathbf{r})\Big)\phi_k(\mathbf{r}) = \epsilon_k \phi_k(\mathbf{r}). \tag{4.3}$$

The density $\rho[v_s](\mathbf{r})$ can be computed using Eq. 2.6. To optimize the Lagrangian in Eq. 4.2, we analytically compute the gradient of L_{KS} with respect to coefficients $\{c_P\}$

$$\frac{\partial L_{KS}}{\partial c_P} = \int \widetilde{\chi}_P(\mathbf{r})\big(\rho[v_s](\mathbf{r}) - \rho_q(\mathbf{r})\big)d\mathbf{r}, \tag{4.4}$$

and the Hessian

$$\frac{\partial^2 L_{KS}}{\partial c_P \partial c_Q} = \int\int \widetilde{\chi}_P(\mathbf{r}) \cdot 4 \sum_i^{occ} \sum_a^{vir} \frac{\phi_i(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r}')}{\epsilon_i - \epsilon_a} \cdot \widetilde{\chi}_Q(\mathbf{r}') d\mathbf{r} d\mathbf{r}'$$
$$= 4 \sum_i^{occ} \sum_a^{vir} \frac{(\phi_i\phi_a|\chi_P)(\phi_i\phi_a|\chi_Q)}{\epsilon_i - \epsilon_a}. \tag{4.5}$$

Note that the magnitude of the gradient in Eq. 4.4 is an indicator for the density matching. With analytic gradient and Hessian, we can use a standard Newton's method to find the optimized potential expansion coefficients $\{c_P\}$. Once the converged KS orbitals $\{\phi_k(\mathbf{r})\}$ are found, the DFT energy for $\rho_q(\mathbf{r})$ can be computed: $E_a[\rho_q] = E_a[\{\phi_k\}]$.

To obtain the exact energy of the same fragment density $\rho_q(\mathbf{r})$, we apply an interacting Wu-Yang potential inversion algorithm. We first expand the potential $v_{ex}(\mathbf{r})$ in Eq. 2.9:

$$v_{ex}(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_P f_P \widetilde{\chi}_P(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_P f_P \int \frac{\chi_P(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \tag{4.6}$$

Then Eq. 2.9 becomes

$$L_{Exact}[\Psi, \{f_P\}] = \langle\Psi|[\frac{1}{2}\sum_k \hat{\mathbf{p}}_k^2 + \sum_{k<l} \frac{1}{\hat{\mathbf{r}}_{kl}}]|\Psi\rangle$$
$$+ \int \Big(v_{nuc}(\mathbf{r}) + \sum_P f_P \widetilde{\chi}_P(\mathbf{r})\Big)(\rho[v_{ex}](\mathbf{r}) - \rho_q(\mathbf{r}))d\mathbf{r}. \tag{4.7}$$

and Ψ is the ground state of the interacting Schrödinger equation

$$[\frac{1}{2}\sum_i \hat{\mathbf{p}}_i^2 + \sum_{i<l} \frac{1}{\hat{\mathbf{r}}_{il}} + v_{nuc}(\hat{\mathbf{r}}) + \sum_P f_P \widetilde{\chi}_P(\hat{\mathbf{r}})]|\Psi\rangle = E|\Psi\rangle. \tag{4.8}$$

The density $\rho[v_{ex}](\mathbf{r})$ can be computed using Eq. 2.11. The optimization of the Lagrangian in Eq. 4.7 can also be performed using a standard Newton's method, as in the non-interacting case. The analytic gradient can be computed as

$$\frac{\partial L_{Exact}}{\partial f_P} = \int \widetilde{\chi}_P(\mathbf{r})\big(\rho[v_{ex}](\mathbf{r}) - \rho_q(\mathbf{r})\big)d\mathbf{r}, \tag{4.9}$$

while the Hessian needs to be computed using the finite difference method. The converged interacting wave function Ψ can then be used to compute the exact energy for the fragment density $\rho_q(\mathbf{r})$: $E_v[\rho_q] = \langle\Psi|\hat{\mathbf{H}}|\Psi\rangle$. Note that although there is no guarantee that the SAH

pair densities are also interacting v-representable, our numerical tests show that one can usually find the corresponding potential $v_{ex}(\mathbf{r})$ in the interacting potential inversion as long as the pair densities are non-interacting v-representable in the given basis set.

With the approximate DFT and exact energies for fragment densities, the MPE energies can be finally computed according to Eq. 2.2 at the desired levels. In this study, due to the high computational cost of interacting potential inversion calculations, we perform MPE calculations up to the second order (MPE2), which means we only correct the DFT energies for one-pair and two-pair densities.

As discussed in Chapter 3, the quality of SAH pair densities is affected by the choice of the parameter α. A small α value results in pair densities that are not localized enough, while a large α value may give rise to sharp changes and nodes in pair densities. Note that although the SAH pair densities are ground-state solutions of Kohn-Sham equations, they can still possess nodes due to the use of finite basis sets. Therefore, we want to choose the maximum α value that still maintains smoothness in pair densities. In this work, we choose $\alpha = 4$ for all tested molecules and show the dependence of MPE energies on the choice of α is weak in a later section.

By comparing Eq. 3.14 and Eq. 4.3, one may realize that the potential $v_s(\mathbf{r})$ in the non-interacting potential inversion for one-pair densities are already pre-defined by the SAH decomposition:

$$v_s(\mathbf{r}) = v_{nuc}(\mathbf{r}) + \sum_P c_P \widetilde{\chi}_P(\mathbf{r}) = v_{nuc}(\mathbf{r}) - 2\alpha \sum_P d_P^i \widetilde{\chi}_P(\mathbf{r}) + \sum_t b_t \widetilde{\chi}_t(\mathbf{r}). \tag{4.10}$$

Thus, the expansion coefficients for one-pair densities $\{c_P\}$ can be directly obtained from

$$c_P = -2\alpha \cdot d_P^i + b_P, \tag{4.11}$$

without doing the potential inversion procedure in Eqs. 4.2-4.5.

4.2.2 Energy estimation for non-v-representable densities

For fragment densities larger than one-pair densities, there is no guarantee that they are also v-representable. Although they are constructed from adding smooth one-pair densities, it may still be difficult to find the corresponding converged potentials in the potential inversion

processes, due to the use of finite basis set. Without the converged potentials $v_s(\mathbf{r})$ and $v_{ex}(\mathbf{r})$, the DFT and FCI energies for the fragment density $E_a[\rho_q]$ and $E_v[\rho_q]$ cannot be computed. For example, in the water molecule, the two-pair density that comes from adding two lone-pair densities is not v-representable in the chosen basis set.

For simplicity, assume this two-pair density comes from adding the first and second pair densities $\rho_q(\mathbf{r}) = \rho_{1+2}(\mathbf{r}) = \rho_1(\mathbf{r}) + \rho_2(\mathbf{r})$. We need to estimate the contribution to the energy correction from this non-v-representable two-pair density in Eq. 2.2: $E^{corr}[\rho_{1+2}] = \Delta E[\rho_{1+2}] - \Delta E[\rho_1] - \Delta E[\rho_2]$, where only $\Delta E[\rho_1]$ and $\Delta E[\rho_2]$ can be accurately computed. A straightforward way is to approximately compute the energy correction term $\Delta E[\rho_{1+2}]$. Although the converged potential for non-interacting potential inversion $v_s(\mathbf{r})$ cannot be found, we can instead search for the potential $\widetilde{v}_s(\mathbf{r})$ that gives the most closely approximate density $\widetilde{\rho}_{1+2}(\mathbf{r})$, which is the potential that gives the smallest gradient (Eq. 4.4) norm in the optimization of the Lagrangian in Eq. 4.2. Using the approximate potential $\widetilde{v}_s(\mathbf{r})$, we can estimate the DFT energy for $\rho_{1+2}(\mathbf{r})$: $E_a[\rho_{1+2}] \approx E_a[\widetilde{\rho}_{1+2}]$. An interacting potential inversion can then be performed exactly for the same approximate density $\widetilde{\rho}_{1+2}(\mathbf{r})$ to obtain its FCI energy $E_v[\widetilde{\rho}_{1+2}]$. Since $\widetilde{\rho}_{1+2}(\mathbf{r})$ is a non-interacting v-representable density, we find in practice the convergence of its interacting potential inversion is normally good. The estimation to $E^{corr}[\rho_{1+2}]$ can thus be calculated as:

$$\begin{aligned} E^{corr}[\rho_{1+2}] &= \Delta E[\rho_{1+2}] - \Delta E[\rho_1] - \Delta E[\rho_2] \approx \Delta E[\widetilde{\rho}_{1+2}] - \Delta E[\rho_1] - \Delta E[\rho_2] \\ &\approx E_v[\widetilde{\rho}_{1+2}] - E_a[\widetilde{\rho}_{1+2}] - \Delta E[\rho_1] - \Delta E[\rho_2]. \end{aligned} \tag{4.12}$$

Nevertheless, the approximation in Eq. 4.12 may still lead to large energy errors when the density error $||\mathbf{W}||$ is large, where

$$W_p = \int \widetilde{\chi}_P(\mathbf{r})\big(\widetilde{\rho}_{1+2}(\mathbf{r}) - \rho_{1+2}(\mathbf{r})\big)d\mathbf{r}, \tag{4.13}$$

because $\widetilde{\rho}_{1+2}(\mathbf{r}) \neq \rho_1(\mathbf{r}) + \rho_2(\mathbf{r})$. To solve this problem, we introduce a "top-down" correction scheme. Once we get the approximate two-pair density $\widetilde{\rho}_{1+2}(\mathbf{r})$, we use this density as the input to perform a new SAH decomposition to generate two new one-pair densities:

$$\widetilde{\rho}_{1+2}(\mathbf{r}) \xrightarrow{\text{SAH}} \widetilde{\rho}_1(\mathbf{r}) + \widetilde{\rho}_2(\mathbf{r}), \tag{4.14}$$

where $\widetilde{\rho}_1(\mathbf{r})$ and $\widetilde{\rho}_2(\mathbf{r})$ resemble the original one-pair densities $\rho_1(\mathbf{r})$ and $\rho_2(\mathbf{r})$. In this way, we rebuild the one-pair densities whose sum equals to $\widetilde{\rho}_{1+2}(\mathbf{r})$ and then estimate $E^{corr}[\rho_{1+2}]$ using:

$$\begin{aligned} E^{corr}[\rho_{1+2}] &= \Delta E[\rho_{1+2}] - \Delta E[\rho_1] - \Delta E[\rho_2] \approx \Delta E[\widetilde{\rho}_{1+2}] - \Delta E[\widetilde{\rho}_1] - \Delta E[\widetilde{\rho}_2] \\ &\approx E_v[\widetilde{\rho}_{1+2}] - E_a[\widetilde{\rho}_{1+2}] - (E_v[\widetilde{\rho}_1] - E_a[\widetilde{\rho}_1]) - (E_v[\widetilde{\rho}_2] - E_a[\widetilde{\rho}_2]). \end{aligned} \tag{4.15}$$

Here, we compute the DFT and FCI energies for new one-pair densities $\widetilde{\rho}_1(\mathbf{r})$ and $\widetilde{\rho}_2(\mathbf{r})$, and use those energies to estimate $E^{corr}[\rho_{1+2}]$. This is a better cancellation of error compared to Eq. 4.12 due to the density match $\widetilde{\rho}_{1+2}(\mathbf{r}) = \widetilde{\rho}_1(\mathbf{r}) + \widetilde{\rho}_2(\mathbf{r})$.

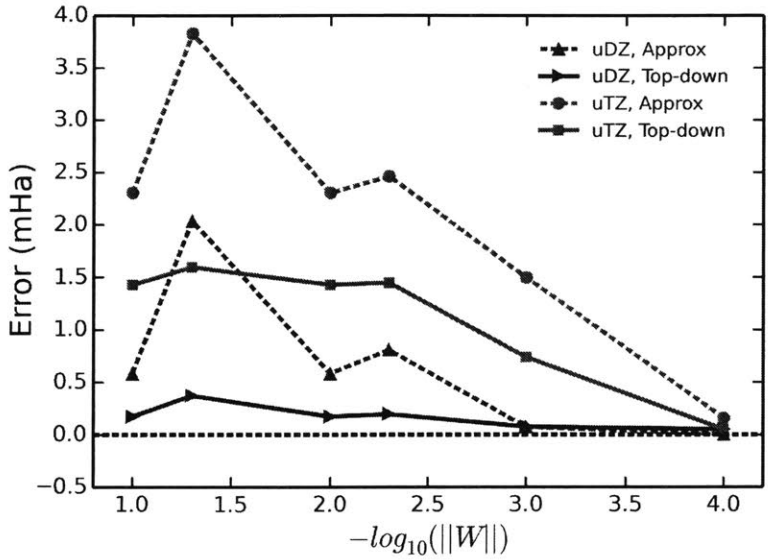


Figure 4-1: The errors of energy estimation for a two-pair fragment density (core + O-H bond) in H_2O.

We conduct a specific numerical test to show the energy errors associated with the estimations made in Eqs. 4.12 and 4.15, as shown in Figure 4-1. The two-pair density of interest is constructed from adding the core pair density and O-H bonding pair density in H_2O, where both non-interacting and interacting potential inversions can converge very well. We choose this two-pair density because its correct DFT and FCI energies are known so that they can serve as the reference values. We manually stop the non-interacting potential inversion process under different thresholds so that the density error is chosen to be: $||\mathbf{W}|| = 10^{-4}, 10^{-3}, 5 \times 10^{-2}, 10^{-2}, 5 \times 10^{-1}, 10^{-1}$ au. Using these approximate densities,

we compute $E^{corr}[\rho_{1+2}]$ according to Eqs. 4.12 (marked as "Approx") and 4.15 (marked as "Top-down") in two different basis sets: uncontracted cc-pVDZ (uDZ) and uncontracted cc-pVTZ (uTZ) [231] in Figure 4-1. One can see when the density error increases, the energy error of $E^{corr}[\rho_{1+2}]$ can be as large as 2 mHa in uDZ basis and almost 4 mHa in uTZ basis for the "Approx" scheme. Applying the "Top-down" correction scheme reduces the energy error to under 0.5 mHa in uDZ basis and approximately 1.5 mHa in uTZ basis. Therefore, the "Top-down" correction scheme reduces the energy errors of the "Approx" scheme by approximately 2-4 folds, but still leads to notable errors. In practice, for some non-v-representable two-pair densities, the non-interacting potential inversion process may only be converged to give $||\mathbf{W}|| = O(10^{-2}$ a.u.), so the energy errors of $E^{corr}[\rho_{1+2}]$ are estimated to be $O(m\text{Hartree})$. We find that molecules with lone pairs or multiple bonds tend to have this numerical instability in MPE calculations. In this work, only the MPE2 results involving CH_2, H_2O and HF are affected by this numerical error.

4.3 Computational details

The ground state geometries of tested small molecules are optimized using the coupled-cluster singles and doubles (CCSD) method [264] in the cc-pVTZ basis set [231] using Q-Chem 4.2 software package [230]. For SAH decomposition and MPE calculations, two basis sets are used for tests: (a) uncontracted cc-pVDZ [231], abbreviated as uDZ; (b) uncontracted cc-pVTZ for non-hydrogen atoms and uncontracted cc-pVDZ for hydrogen atoms, abbreviated as uTZ. The input electron density for all molecules is obtained through Hartree-Fock (HF) calculations in uDZ and uTZ basis sets. In this study, we only use the exact exchange (EXX) functional as our approximate density functional, and perform all MPE calculations starting from EXX results. Coupled-cluster singles and doubles with perturbative triples (CCSD(T)) [36] energies are computed using Q-Chem 4.2 package as the reference for thermochemistry results. Density matrix renormalization group (DMRG) [178] energies are calculated using PySCF package [265] and BLOCK code [179, 266] as the reference for hydrogen ring bond breaking results.

As discussed in Chapter 3, the potential inversion and optimized effective potential (OEP) techniques are numerically unstable in finite basis sets [233, 234]. Several solutions exist to avoid ill-conditioning problems in potential inversions [235, 236, 237, 82]. Here, we

follow the work of Görling and co-workers [235] to balance the orbital and auxiliary basis sets. This leads to our use of uncontracted orbital basis sets in this work. Meanwhile, the corresponding auxiliary basis set $\{\chi_P(\mathbf{r})\}$ for expanding potentials is carefully constructed out of the orbital basis set by removing some of the most compact and diffuse functions. The detailed description of balanced orbital and auxiliary basis sets can be found in the Appendix. VMD [238] and Molden [239] are used to generate fragment density pictures.

4.4 Results

4.4.1 Thermochemistry

We first test the performance of MPE on the thermochemistry of molecules. DFT is known to perform very well for thermochemistry due to the effective description of dynamic correlations [23]. Hybrid and double-hybrid functionals [49, 26, 53] are especially efficient, whose parameters are normally fitted to reproduce experimental thermochemistry data. However, even for thermochemistry, approximate functionals cannot achieve the chemical accuracy of 1 kcal/mol for random chemical reactions. Thus, systematic improvement of density functional calculations is still needed for the accurate description of thermochemistry.

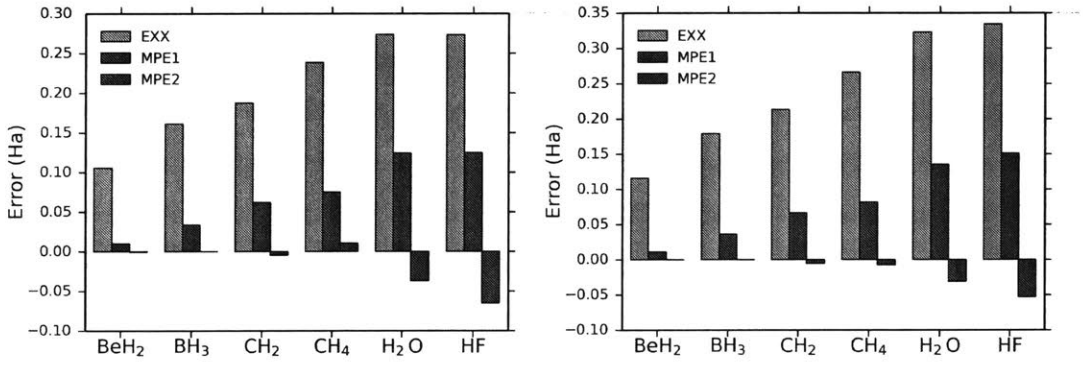


Figure 4-2: Errors of MPE energies for a series of small molecules using CCSD(T) results as the reference. (a) uDZ basis set. (b) uTZ basis set.

We apply MPE to a series of small molecules on their equilibrium geometries, including BeH_2, BH_3, CH_2, CH_4, H_2O and HF, and show the errors of total molecular energies in Figure 4-2. As can be seen, MPE at the first order (MPE1) reduces the errors of EXX

significantly – more than 50% for all tested molecules in both basis sets, meaning that the correlation energy inside each electron pair is very large. However, MPE1 results still possess considerable errors, indicating that correcting the energy errors at only one-pair level is not enough. MPE at the second order (MPE2) further reduces the errors and gives very accurate total energies for BeH_2, BH_3, CH_2 and CH_4. Although MPE2 causes over correlations for H_2O and HF, which may result from the numerical instability in the potential inversions and low order of MPE expansion, it is still a great improvement over EXX and MPE1. This result suggests that adding in correct pair-pair interactions is crucial for the accurate description of dynamic correlations in molecules. More importantly, Figure 4-2 demonstrates that when we apply successive higher order corrections, the MPE total energies become more accurate and converge quickly towards the highly accurate results.

Comparing Figure 4-2a and Figure 4-2b, one may notice that the MPE results are more accurate in the larger uTZ basis set despite the larger EXX errors. As shown in Eq. 4.4 and Eq. 4.9, the quality of density matching in the potential inversions are determined by the size of the auxiliary basis set $\{\chi_P(\mathbf{r})\}$. Therefore, the superior performance of MPE in the larger basis set is probably due to the better density matching in the potential inversions, which implies MPE would be even more accurate when approaching the complete basis set limit.

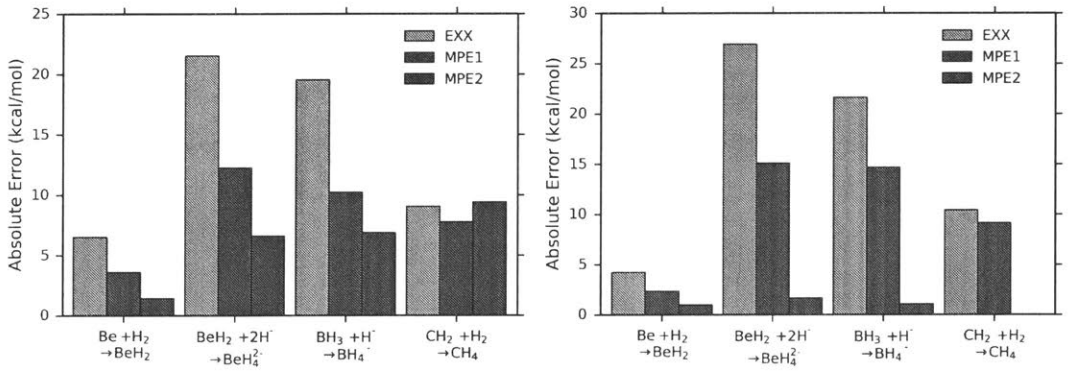


Figure 4-3: Absolute errors of MPE reaction energies using CCSD(T) results as the reference. (a) uDZ basis set. (b) uTZ basis set.

We then investigate the performance of MPE for four chemical reactions, as shown in Figure 4-3. In both basis sets, MPE1 improves the description of reaction energies over

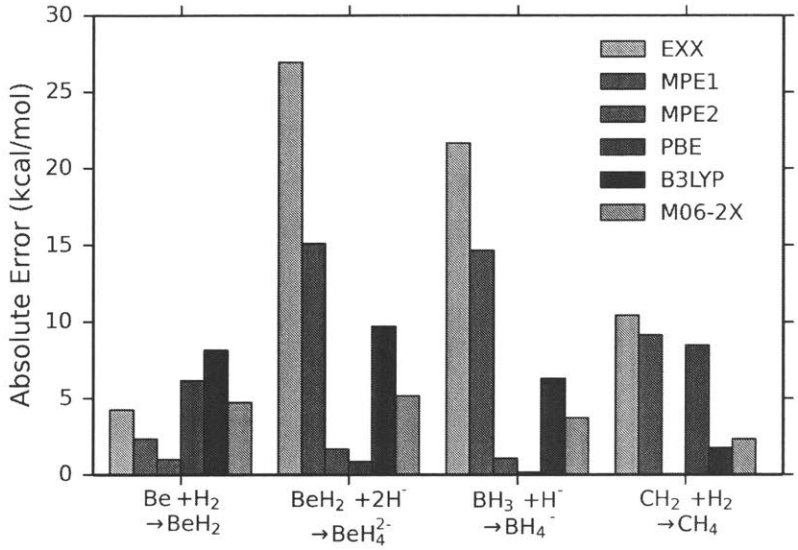


Figure 4-4: Absolute errors of MPE reaction energies compared to several popular density functionals in the uTZ basis set. CCSD(T) results are used as the reference. Note that all MPE calculations use EXX as the starting functional.

EXX for all chemical reactions considered. In the uDZ basis set (Figure 4-3a), MPE2 performs better than MPE1 in the first three reactions but worse in the last one. Although MPE provides successive improvement for the first three reactions, the absolute error of MPE2 reaction energy can still be larger than 5 kcal/mol. However, in the larger uTZ basis set (Figure 4-3b), the MPE results get much better. There is clearly systematic improvement of MPE reaction energies for all four chemical reactions. Meanwhile, the absolute errors of MPE2 reaction energies are small in the uTZ basis set, all of which are below 2 kcal/mol. This result again shows the importance of accurate density matching in the potential inversions – the larger basis set results in better performance of MPE.

In Figure 4-4, we further compare the performance of MPE with several popular density functionals on the same four reactions in the uTZ basis set. The considered functionals involve PBE (GGA) [43], B3LYP (hybrid GGA) [49] and M06-2X (hybrid meta-GGA) [26]. One may notice that none of these three functionals is able to predict all reaction energies within 2 kcal/mol of error like MPE2. For example, PBE is the best functional for describing the second and third reactions, but it poses absolute errors of more than 5 kcal/mol for the other two reactions. Overall, Figures 4-3 and 4-4 demonstrate that, even with the worst performer EXX functional as the starting point, MPE is able to systematically correct the

approximate functional and provides an accurate description of thermochemistry at a low second order of expansion.

The thermochemistry of hydrogen molecular chains presents an interesting challenge for density functional approximations. In spite of the simple model, commonly-used GGA and hybrid functionals have been shown to fail in accurately describing the decomposition of hydrogen chains into hydrogen molecules [267, 268]. The failure is attributed to the inaccurate characterization of the delocalized electron distribution across more than one H_2 unit, which is caused by the self-interaction error (delocalization error) present in these approximate functionals. We therefore test the performance of MPE on the decomposition of a hydrogen chain H_{2n} into n hydrogen molecules, whose reaction energy can be calculated as:

$$\Delta E = nE(H_2) - E(H_{2n}), \tag{4.16}$$

where n is the number of H_2 units in the hydrogen chain. The bond lengths between two adjacent H atoms in the hydrogen chains and hydrogen molecules are set to be 0.75 Å. The static correlation becomes another source of error for the tested methods due to the equal spacing between H atoms in the H_{2n} chain, but it is not dominant because of the small bond length.

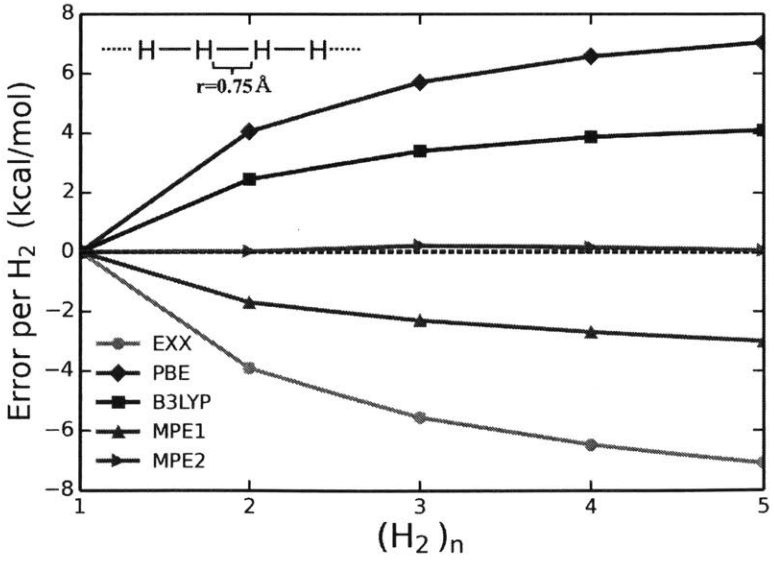


Figure 4-5: Errors of the hydrogen chain reaction energies per H_2 unit in the uDZ basis set. CCSD(T) results are used as the reference.

The errors of the hydrogen chain decomposition energies per H_2 unit compared to CCSD(T) is shown in Figure 4-5. The popular GGA functional PBE is shown to overestimate reaction energies severely: about 7 kcal/mol of error per H_2 unit for H_{10} chain decomposition, indicating the large delocalization error in the PBE functional. On the other hand, the EXX functional significantly underestimates the reaction energies, with the similar magnitude of errors as PBE. This behavior is associated with the so-called localization error in literature, due to the neglect of electron correlation in the EXX functional. The hybrid functional B3LYP reduces the errors by one third compared to PBE because of the mixing of EXX, but still poses substantial errors. Furthermore, the errors of all three functionals increase as the hydrogen chain length becomes longer, which suggests the delocalization/localization errors are more severe in more extensive electron density distributions.

As seen in Figure 4-5, MPE1 provides a significant improvement over EXX, where the errors of reaction energies are already smaller than B3LYP by including the correct correlation for each electron pair in the hydrogen chain. The higher order MPE2 produces even better results: errors for different lengths of hydrogen chains are all reduced to almost zero, which indicates that including one-pair and pair-pair electron correlations is sufficient to remove all localization errors present in the EXX functional for hydrogen chains. In addition, this systematic improvement does not deteriorate as the chain length becomes longer, showing the ability of MPE to deal with systems with different electron density distributions.

In all test cases for thermochemistry, we use the Hartree-Fock (HF) density as the input to perform MPE calculations. Due to this approximate total electron density, MPE will not converge to the true exact (FCI) energy. Instead, MPE energies converge to the FCI energy constrained to the given HF density, which equals to inverting the full interacting Schrödinger equation with the density constraint. The HF density is known to be a good approximation of electron density for equilibrium molecular structures [269], so we expect it causes small energy errors. Our results above also show the effect of using HF density is small for the description of thermochemistry.

4.4.2 Strong correlation

To test the performance of MPE on strongly correlated molecular systems, we study the symmetric bond breaking process in hydrogen rings using MPE. Hydrogen rings have recently

emerged as a benchmark for new electronic structure methods [270, 271], as the strength of correlation can be tuned from weak to strong by changing the hydrogen atom spacing. Despite the simple model, the strong electron correlations present in stretched hydrogen rings require the correlation methods to be capable of treating full Coulomb interactions. Meanwhile, hydrogen rings in a realistic basis set have multiple orbitals per site as well as long-range interactions, making them more complicated compared to the one-dimensional (1D) Hubbard model [140]. We have shown that MPE can deal with strong correlations in 1D Hubbard model very well [139], so it is interesting to investigate its performance on hydrogen rings, which reflects the challenge of real molecular systems.

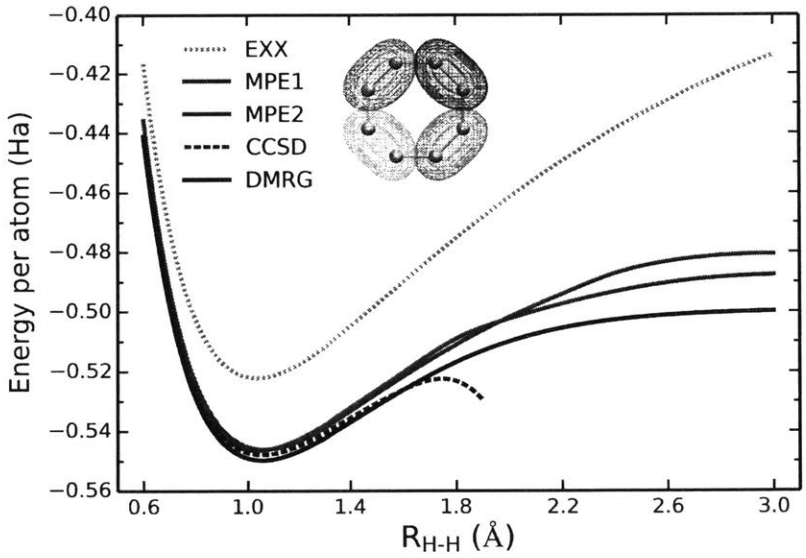


Figure 4-6: Energies per H atom in the H_8 ring in the uDZ basis set. SAH decomposed pair densities are shown as an inset.

We first show the energies per H atom in the H_8 ring in the uDZ basis set in Figure 4-6. The DMRG result is presented here as a highly accurate answer for comparison. The EXX functional is shown to predict much higher energies compared to DMRG around the equilibrium distance, indicating it is incapable of describing dynamic correlations in the H_8 ring. When the H-H distance gets longer, the strong (static) correlation becomes more dominant, leading to even worse performance of EXX functional. This is expected for all single-determinantal approximate KS-DFT methods. Adding in correct one-pair correlation in MPE1 results in a great improvement over EXX. The equilibrium-geometry result

of MPE1 is already accurate compared to DMRG. More importantly, in the bond-breaking region, MPE1 reduces the EXX errors substantially because the use of FCI brings in correct multi-determinantal interactions at the two-electron level. This correction cannot be achieved using traditional single-reference correlation methods such as CCSD. As can be seen in Figure 4-6, CCSD describes the dynamic correlations very well but collapses when the H-H distance is long [271]. Despite the much better accuracy than EXX, it is clear that MPE1 is still insufficient for accurate estimation of strong correlations since the error per atom is still as high as 0.02 Ha at $R_{H-H} = 3.0$ Å.

Including pair-pair electron correlations leads to further improvement in MPE2 compared to MPE1. MPE2 is still accurate around the equilibrium distance and further reduces the MPE1 errors at longer H-H distances. Nevertheless, unlike in thermochemistry cases, even MPE2 cannot treat strong correlations in the stretched H_8 ring accurately. This result can be attributed to two reasons. First, the second order of expansion for MPE may not be enough for this problem. Higher order of expansions, where six or more electron correlations are correctly computed, may be needed. Second, a more important reason might be that we employ the HF density as the input for MPE calculations here. As already shown in Figure 4-6, EXX, which is the counterpart method of HF in DFT, is a very poor solution for the strong correlation problems. Using bad input electron density from HF may cause MPE to finally converge to the inaccurate result, i.e., it is possible in this case even MPE4 is not satisfactory. Therefore, to achieve a correct description of strong correlations, we may want to have better input electron density. Such density may be obtained through multi-configurational SCF calculations, or more ideally, from the optimization of self-consistent MPE calculations. We will not discuss the optimization of input density here and leave it for the future work.

We also apply MPE to the H_6 ring in the uDZ basis set to test its performance for a hydrogen ring with different size, as demonstrated in Figure 4-7. MPE1 and MPE2 give similar accuracy at long H-H distances compared to Figure 4-7: MPE1 provides a significant improvement over EXX and MPE2 further reduces the errors of MPE1, while both MPE1 and MPE2 pose errors of more than 0.01 Ha per atom at $R_{H-H} = 3.2$ Å. Around the equilibrium distance, however, the performance of MPE is not as satisfactory as in the H_8 ring. Both MPE methods overcorrelates for about 0.01 Ha per atom compared to DMRG and MPE2 provides no improvement over MPE1. This result is surprising considering there

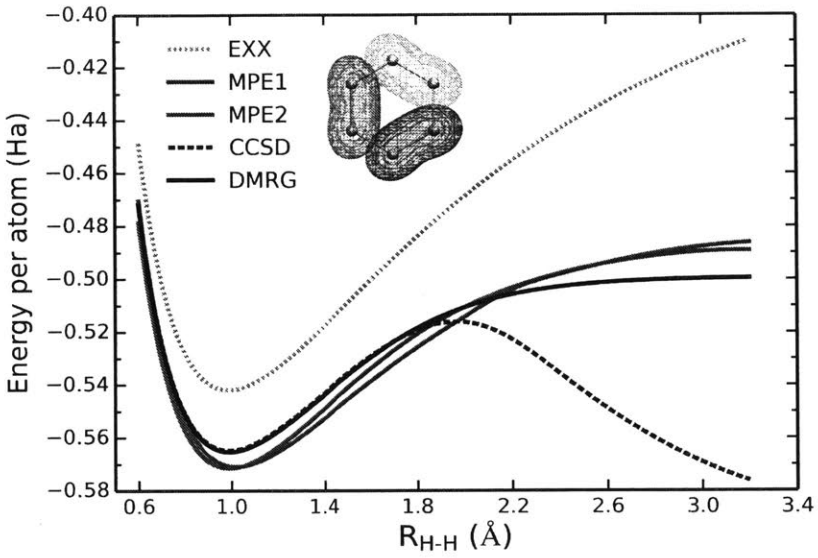


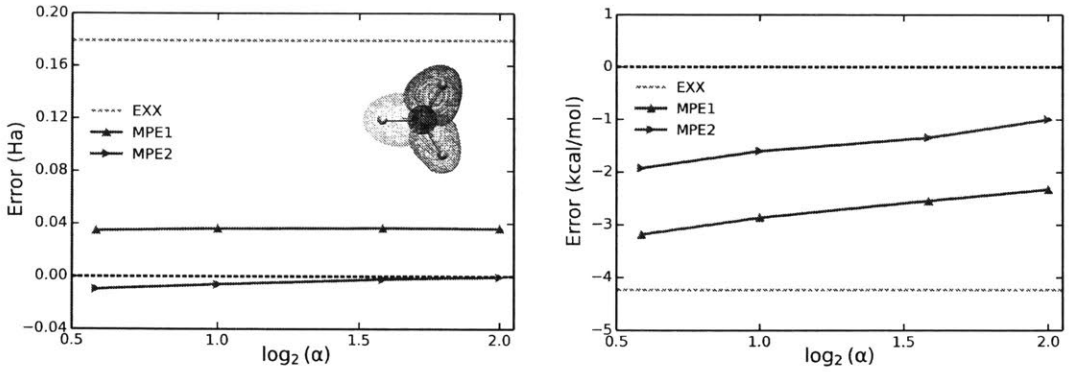
Figure 4-7: Energies per H atom in the H_6 ring in the uDZ basis set. SAH decomposed pair densities are shown as an inset.

are only six electrons in this system, meaning that MPE3 is the "exact" answer. We suspect such behavior is related to the aromaticity of the system: the H_6 ring can be viewed as an aromatic ring while the H_8 ring is anti-aromatic. In aromatic systems, specific order of MPE (for example, MPE3) and better density decomposition may be needed to achieve high accuracy, which is also found in the previous MPE study of benzene dimer in the PPP model [145] in Chapter 2.

4.4.3 Dependence on localization parameter α

All MPE results shown above are based on pair densities generated by the SAH decomposition using the localization parameter $\alpha = 4$ for Eq. 3.14. We choose $\alpha = 4$ because it is the maximum value that still maintains smoothness in pair densities so that the ill-conditioning problems in potential inversions can be mostly avoided. Meanwhile, the resulting pair densities are localized enough to make sure MPE energies converge quickly to the accurate result. Here, we test the performance of MPE with different α values ($\alpha = 1.5, 2, 3, 4$) to show its dependence on the parameter α.

We fist show the dependence of MPE on the localization parameter α for the total molecular energy of BH_3 in Figure 4-8a. For all four α values tested, the MPE1 energy is



(a) (b)

Figure 4-8: The dependence of MPE energies on the localization parameter α in the uTZ basis set. CCSD(T) results are used as the reference. (a) Total energy of BH_3. (b) Reaction energy of Be + H_2 $\rightarrow$ BeH_2.

quite stable, while the MPE2 energy becomes slightly better when larger α value is used. Similar results can be observed in Figure 4-8b, where we test the dependence of MPE on α for the reaction energy of Be + H_2 $\rightarrow$ BeH_2. Both MPE1 and MPE2 predict better reaction energies when larger α values are used. However, the difference of MPE2 reaction energy for using $\alpha = 1.5$ and $\alpha = 4$ is just 1 kcal/mol. Overall speaking, we find that more localized SAH decomposition (larger α value) leads to better accuracy of MPE, and the dependence of MPE on the parameter α is not strong. Thus, we believe $\alpha = 4$ is a suitable choice for this study, and $\alpha \geq 4$ should be used in the future MPE study as long as the potential inversions remain numerically stable.

4.5 Conclusions

In this chapter, we extend the many-pair expansion to molecular systems with *ab initio* Hamiltonians. We employ the self-attractive Hartree decomposition to generate localized and v-representable pair densities as the basic fragment densities in molecular MPE. To perform numerically stable potential inversions, we introduce auxiliary basis sets to represent the external potentials and construct the orbital and potential basis sets in a balanced fashion. Our numerical study shows that MPE predicts accurate molecular and reaction energies for a few small molecules and hydrogen chains, even with the low-level EXX functional and a low second order of expansion. The description of thermochemistry is systematically

improved as successive higher order MPE corrections are applied. Meanwhile, through the hydrogen chain case, we demonstrate the potential of MPE to serve as a method to systemically remove the many-electron self-interaction error in approximate density functionals. The study of bond breaking processes in hydrogen rings exhibits the ability of MPE to deal with strong correlations, but also points out the need for better input density or density decomposition in MPE. In conclusion, this work shows that MPE provides a promising framework to systematically improve density functional calculations of molecules. We will further discuss the direction of future MPE work in Chapter 7.

4.6 Acknowledgements

This work is carried out together with Dr. Piotr de Silva.

Chapter 5

Condensed phase simulation of host-guest OLEDs

5.1 Introduction

In the previous chapters, we have presented a method to systematically improve the accuracy of density functional calculations. While our work to extend the many-pair expansion to more realistic systems is still ongoing, we also want to point out that conventional DFT methods can be powerful tools to model electronic properties of complex chemical systems as long as they are utilized properly. In this and next chapters, we explore this aspect of DFT simulations by using organic light-emitting diodes (OLEDs) [27] as an example. We choose OLEDs as the model system because of two reasons. First, to understand the charge and energy transfer mechanisms in OLEDs, both ground-state (e.g., ionization potential (IP) and electron affinity (EA)) and excited-state (e.g., T_1 and S_1 excitation energies) energetic properties need to be accurately calculated, which posts a challenge to DFT methods. Second, OLEDs are disordered condensed phase systems, whose electronic properties are affected significantly by the molecular environment. This requires one to correctly incorporate environment effects into DFT simulations using the methods we mentioned in Chapter 1. In this chapter, we carry out condensed phase simulations to investigate the charge recombination and energy transfer processes in host-guest phosphorescent OLEDs [272].

Organic light-emitting diodes (OLEDs) have attracted significant attention for broad applications in displays and lighting because of their high electroluminescence (EL) efficiency,

flexibility and low manufacturing cost [273, 274, 275, 276, 277]. In order to improve the EL efficiency, various fluorescent and phosphorescent materials have been introduced as OLED emitters. Because the ratio of singlet and triplet excitons under electrical excitation is 1:3 due to spin statistics [278, 279, 280], the internal quantum efficiency (IQE) of traditional fluorescent OLEDs is limited to 25%. Phosphorescent OLEDs (PhOLEDs), on the other hand, can achieve 100% IQE by harvesting both singlet and triplet excitons through strong spin-orbit coupling [281, 282, 283, 284] and thus obtain high external quantum efficiency (EQE) over 20% [285, 286, 287, 288]. Recently, a new type of fluorescent OLEDs, thermally activated delayed fluorescence (TADF), is developed to also achieve 100% IQE through efficient reversed intersystem crossing from the T_1 state to S_1 state [289, 290].

No matter what kind of emitter is adopted, the host-guest system is necessary in OLED devices to achieve high external quantum efficiency. In host-guest systems, emitters (guests) are dispersed homogeneously into host materials in order to minimize triplet-triplet annihilation and exciton quenching [291, 292]. As shown by numerous experiments [293, 294, 295], the electroluminescence efficiency of OLEDs is affected significantly by different choices of host-guest combinations because guest exciton formation relies on efficient charge and energy transfers from the host. Other important properties, such as the color purity [296] and exciton lifetime [297] in OLED devices, also rely on tailored host-guest design. Therefore, understanding host-guest interactions and how they affect energy and charge transfers is crucial for improving performance of OLED devices.

However, the mechanism of guest exciton formation in host-guest systems is not well understood and generally two hypotheses have been proposed. First, the holes and electrons recombine primarily on the host, and then the excitation energy is transferred from the host to the guest. This process happens through Förster resonance energy transfer (FRET) [298] for singlet excitons and Dexter energy transfer (DET) [299] for both singlet and triplet excitons. Several new design strategies based on this mechanism are suggested, such as introducing triplet sensitizers [300], exciplex [301] or TADF molecules [302] to assist energy transfer to the emitters. Second, several other studies reported that charge trapping at guest molecules is the main cause for guest exciton formation [303, 304]. This mechanism assumes the guest can capture a charge carrier as a trap site. Such charge carrier can then recombine with a charge carrier of opposite sign to form an exciton. However, it should be pointed out even if some guests may trap charge and form excitons directly, effective quenching of

host excitons is still necessary to avoid host emission [274]. Therefore, it is believed both mechanisms should contribute to the guest emission in a certain ratio [305, 306].

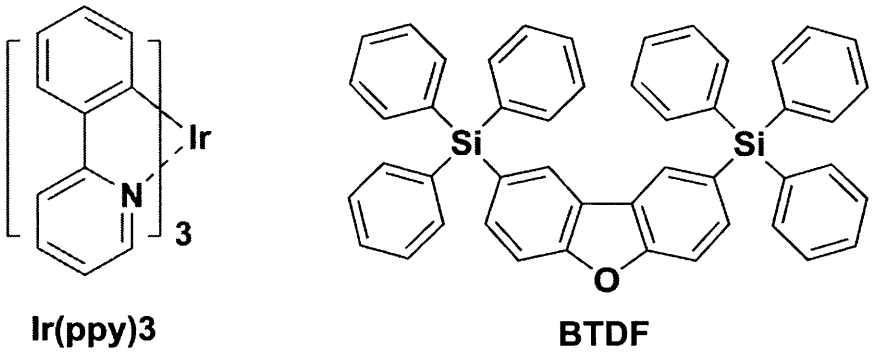


Figure 5-1: Molecular structures of $Ir(ppy)_3$ and BTDF.

In this chapter, we present a theoretical investigation on a model PhOLED emission layer to understand these two mechanisms, in which the organometallic complex *fac*-tris(2-phenylpyridine) iridium (*fac*-$Ir(ppy)_3$) and 2,8-bis(triphenylsilyl) dibenzofuran (BTDF) are chosen as the guest and host, as shown in Figure 5-1. $Ir(ppy)_3$ is a well-known green phosphorescent emitter reported to achieve a maximum EQE over 20% [285, 286, 307, 308], which makes it a good candidate for full color displays. BTDF is an electron-conducting host designed for hole-conducting deep-blue emitters where the maximum EQE of such OLEDs are above 17% [309]. Andrienko and coworkers did a number of computational studies on the morphology and charge transfer properties of BTDF-hosted PhOLEDs on various length scales [310, 311]. While several theoretical methods exist for treating charge transfer problems in organic optoelectronics [312, 313], we use a combined molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) approach [314, 315, 316]. Particularly, we are interested in two problems: (1) how electrons and holes recombine from the host to the guest; (2) assuming a host exciton is formed, how such host exciton is quenched for guest emission. Understanding the charge recombination and host exciton quenching mechanisms would be of great importance for improving formation and utilization of excitons for guest emission and thus lead to better host-guest design.

5.2 Mechanism of charge recombination on the guest

5.2.1 Computational methods

We begin our study on the $Ir(ppy)_3$-BTDF system with a molecular dynamics (MD) simulation, where the simulation box contains 15 $Ir(ppy)_3$ and 250 BTDF molecules that are treated classically (Figure 5-2, left panel). We set the guest to host mass ratio to be approximately 6% to match the experimental values of $Ir(ppy)_3$-doped systems (3%-10%) [285, 308, 307]. From the MD trajectory, we extract snapshots and then perform a large number of polarizable QM/MM (QM/MMpol) single-point calculations on the cation, anion and neutral states of the guest and the host, in order to obtain IPs and EAs of both species. For each snapshot, a host or a guest molecule is chosen as the QM region (Figure 5-2, right panel), while all the other molecules are described by MM force fields. In addition, as we find the hole transfer from the host to the guest is thermodynamically downhill while the electron transfer is mostly uphill, we also want to investigate the electron transfer from the host to the positively-charged guest and the formation of excitons on the guest. Thus, we also perform excited-state QM/MMpol calculations. The details of MD and QM/MM simulations are described below.

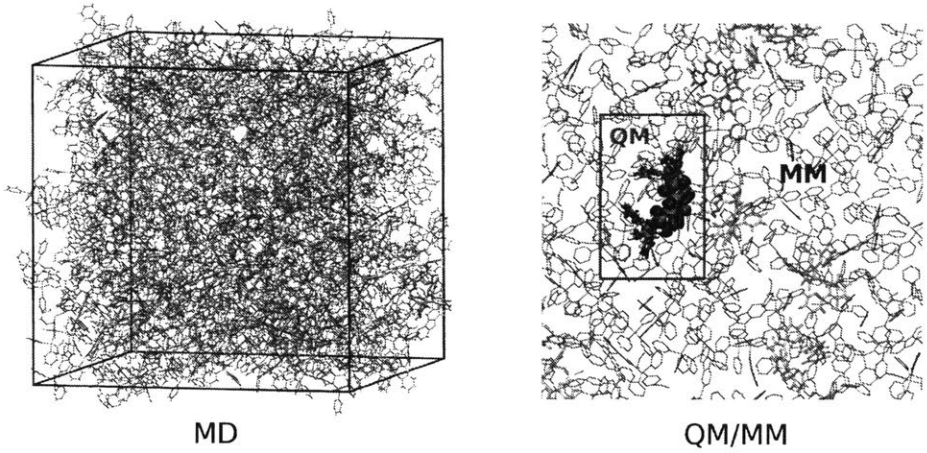


Figure 5-2: Left: Disordered cell of the $Ir(ppy)_3$(green)-BTDF(gray) system generated by MD simulation. Right: Illustration of QM/MM simulation.

MD Simulation. To start the MD simulation, the host and guest molecules are inserted randomly into a simulation box as the initial configuration using PACKMOL package [317].

We anneal the system from 0K to 500K during 2ns and then simulate the system at 500K in an NPT ensemble for another 2ns to make sure the equilibrium density is reached. The host-guest system is equilibrated at the high temperature (500K) so that the molecules become uncorrelated in space and well-sampled amorphous structures can be obtained. The system is then cooled to the room temperature (300K) during 1ns, after which a 3ns simulation in NPT ensemble at 300K is performed. The final 2ns of the NPT dynamics is sampled at 0.2ns intervals to obtain 10 snapshots for QM/MM calculations. MD simulations are performed with a time step of 0.001ps, Berendsen pressure coupling [318] (reference pressure 1 Bar, time constant of τ_P=1ps, and compressibility 4.5×10^{-5} Bar^{-1}), and the grid-based particle mesh Ewald technique [319] for electrostatic interactions. All MD simulations are carried out using GROMACS 4.5.5 package [320] and OPLS force field [321, 309].

To more accurately describe the charge distributions of both host and guest in the force field, we carry out DFT calculations to obtain the CHELPG (Charges from the Electrostatic Potential on a Grid based method) charges [322], which are used to re-parametrize the OPLS charges. The geometries are first optimized using B3LYP functional [49] and 6-311G** basis set [323], while LANL2DZ effective core potential basis set is used for iridium [324]. Then single-point calculations are performed at the same level of theory to obtain CHELPG charges [325].

QM/MM Calculations. All of the QM/MM calculations are done using the CHARMM [326]-QChem [230] interface [327]. All density functional theory (DFT) calculations for QM regions are performed with QChem 4.1 package using the PBE0 functional [328] and 6-31G* basis set [323] for BTDF and LANL2DZ [324] effective core potential basis set for $Ir(ppy)_3$. All of the excited state calculations are done using a restricted open-shell Kohn-Sham (ROKS) approach [59]. ROKS calculates the lowest singlet excited state (S_1) energy by optimizing the KS orbitals to minimize a linear combination of single-determinant energies.

The QM/MM method must take the polarizable organic semiconductor environment into account apart from the electrostatic effects. Therefore, we introduce fictitious "drude" charges in our QM/MMpol method that are harmonically attached to MM charges [101]. These classical drude particles are allowed to respond instantaneously to the electric fields and correspondingly move to their local energy minima positions. The induced polarization

can thus be simulated through such simple treatment. The drude charge parameters are fitted to reproduce the polarizability of the host and guest computed employing DFT.

With QM/MMpol computed energies, we calculate IPs and EAs of the host and guest using the ΔSCF method:

$$IP = E_+ - E_0 \tag{5.1}$$

$$EA = E_0 - E_- \tag{5.2}$$

where E_+, E_-, E_0 are the energies for cation, anion and neutral states.

We perform QM/MMpol calculations on all 250 host molecules and 15 guest molecules from the MD simulation box. The QM region is either a host or a guest molecule. To make the sampling representative enough, we perform calculations for 2 snapshots of 250 host molecules and 10 snapshots of 15 guest molecules.

Reaction Coordinate. To understand the charge recombination in a host-guest system, it is significant to know the hole and electron energy levels of the host in different molecular environment, i.e., bulk phase and near the guest. Since the guest is doped into the host at a small ratio, it can be assumed that the charges are first transferred to the host which is far away from the guest (similar to bulk phase) and then to the host near the guest and finally to the guest. Thus, we need a definition for the "distance" between the host and guest. However, distance between molecules is poorly defined in such system because the molecules are about the similar size (~1 nm) as the space between their centers of mass (1-5 nm). Besides, as there are many guest molecules in the simulation box, it is ambiguous to determine which host-guest distance to use for a specific host.

In this study, we introduce a new reaction coordinate — the "host-guest interaction energy", to evaluate the influence of the guest on the host energy levels. Since the electrostatic energy between two molecules is inversely proportional to their distance, this energy can be used to indirectly measure the distance. To avoid the ambiguity of choosing specific host-guest electrostatic energy, we evaluate the electrostatic effects of all guests on one host by calculating its IP/EA under two different electrostatic environments. Therefore, we define the host-guest interaction energy as the difference of two sets of the host IP/EA. The three steps to obtain the host-guest interaction energy are:

1. Calculate the IP/EA (marked as IP(on)/EA(on)) of the host molecules.

2. Turn off all MM and drude charges on all guests, then re-calculate the IP/EA (IP(off)/EA(off)) of the same set of host.

3. The host-guest interaction energy is given by:

$$|\Delta IP| = |IP(off) - IP(on)| \tag{5.3}$$

$$|\Delta EA| = |EA(off) - EA(on)| \tag{5.4}$$

Using the difference of IP or EA values due to the existence/absence of the guest charges, we find a better way to describe the electrostatic and polarizable effects on the host energetics caused by the guest.

5.2.2 Simulation Results

We first show electron affinities and ionization potentials of the host as a function of the host-guest interaction energy in Figures 5-3 and 5-4. Eq. 5.4 was employed to compute the host-guest interaction energy as the reaction coordinate for EA, while for IP we used Eq. 5.3. In experiments, $-$IP and $-$EA of solid materials are usually measured relative to the reference energy levels using ultraviolet photoemission spectroscopy (UPS) and inverse photoemission spectroscopy (IPES) [329, 330]. In some literatures, $-$IP and $-$EA are called the highest occupied molecular orbital (HOMO)/lowest unoccupied molecular orbital (LUMO) energies, though which is not strictly accurate. Here, we plot $-$IP and $-$EA to be consistent with the experiments.

Clearly there is much scatter in the data (Figures 5-3 and 5-4), although there is also a trend. To be quantitative about the scatter and trend of EA as a function of the distance from the guest, we need to use the data to estimate both the average value of EA at a given point on the reaction coordinate and also its standard deviation. For this purpose we use the kernel density estimation (KDE), a statistical tool for reconstructing probability distributions from scattered data [331, 332]. We use a small Gaussian distribution (the kernel) centered on each data point to simulate its probability. The overall probability distribution of EA is thus given by Eq. 5.5, where $|\Delta EA|_i$ is each data point's value of $|\Delta EA|$, N is the number of data points and α is the coefficient of the Gaussian distribution.

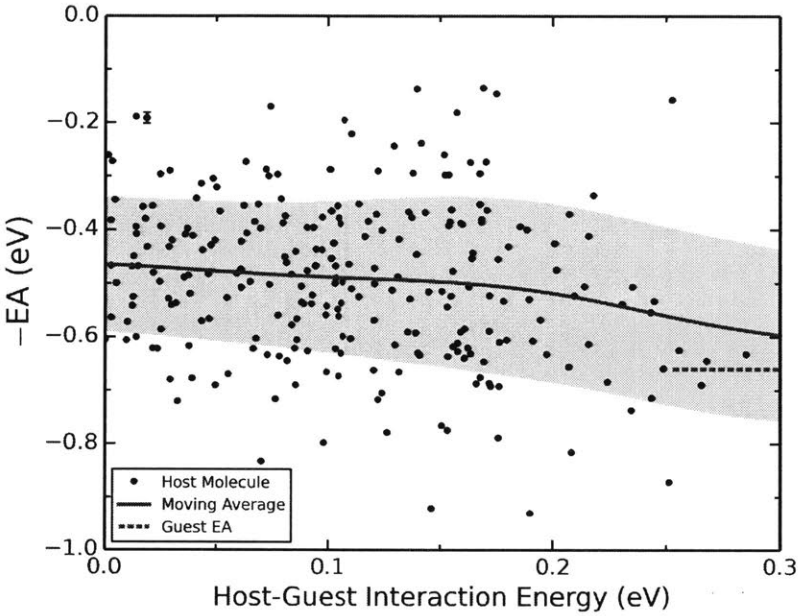


Figure 5-3: Electron affinities of the host as a function of the host-guest interaction energy $|\Delta EA|$. The yellow region covers the moving average plus/minus the moving standard deviation ($\overline{EA} \pm \sigma_{EA}$).

We choose $\alpha = 300$ so that the probability estimation is close to the true probability while the probability curve is kept smooth. The average and the standard deviation of EA are given by Eq. 5.6-5.7, where EA_i is each point's value of EA. For IP, the same equations are used but EA data are replaced by IP data.

$$\rho(|\Delta EA|) = \frac{1}{N}\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} e^{-\alpha(|\Delta EA|-|\Delta EA|_i)^2} \tag{5.5}$$

$$\overline{EA}(|\Delta EA|) = \frac{\frac{1}{N}\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} EA_i \cdot e^{-\alpha(|\Delta EA|-|\Delta EA|_i)^2}}{\rho(|\Delta EA|)} \tag{5.6}$$

$$\sigma(|\Delta EA|) = \sqrt{\frac{\frac{1}{N}\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} (EA_i - \overline{EA}(|\Delta EA|))^2 \cdot e^{-\alpha(|\Delta EA|-|\Delta EA|_i)^2}}{\rho(|\Delta EA|)}} \tag{5.7}$$

Through this KDE procedure, we obtain a smooth moving average curve, which is very useful to describe the IP/EA trend at different distance. Meanwhile, the standard deviation provides a quantitative view of the broadness of the data distribution. It can be seen in Figure 5-3 that the host $-\overline{EA}$ becomes a little lower as the host-guest interaction energy

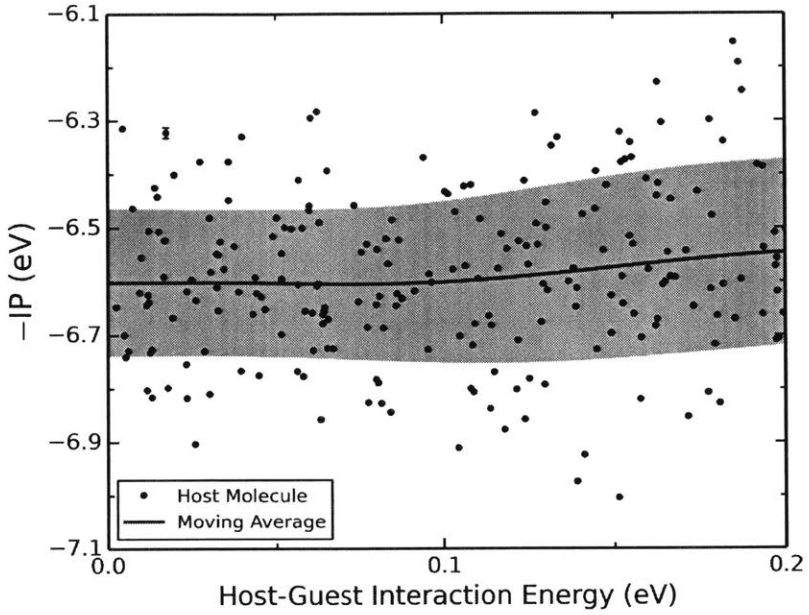


Figure 5-4: Ionization potentials of the host as a function of the host-guest interaction energy $|\Delta IP|$. The green region covers the moving average plus/minus the moving standard deviation ($\overline{IP} \pm \sigma_{IP}$). The guest IP (5.10 eV) is not shown as it is much higher than the host IP.

increases. Note that when the host-guest interaction energy is large, the host molecule is close to a guest. The fact that the change of moving average is not much indicates that the host-guest distance is not a key factor for changing host EA values. A recent QM/MM study of host-guest OLEDs suggests this is because any trend in energy levels is suppressed by the large energetic disorder in OLEDs [333]. In addition, as shown from the scatter plot and the standard deviation (σ_{EA}), the distribution of the host EA is broader as the host gets nearer the guest. This change has an effect on the electron transfer from the host to the guest. When the host is far from the guest, $-\overline{EA} \pm \sigma_{EA}$ is higher than the guest $-EA$ (-0.66 eV), which suggests that the electron transfer would mostly be thermodynamically downhill. However, as the host gets closer to the guest, the distribution of the host EA is approaching the guest EA. Finally, most of the host $-EA$ data are even lower than the guest $-EA$, causing the electron transfer to be much less efficient because the electron transfer becomes an uphill process.

Similar to electron affinities, ionization potentials of the host also show a wider distribution as the host-guest interaction increases, see Figure 5-4. However, this change does

Functional/Basis	6-31G*	6-31+G*	6-311G*
PBE0	6.63/0.87	6.82/0.96	6.81/0.88
B3LYP	6.44/0.54	6.69/0.98	6.68/0.89
PBE	6.08/0.89	6.33/1.29	6.31/1.22

Table 5.1: Basis set and functional effects on BTDF IP/EA values. All numbers are in eV.

Functional/Basis	LANL2DZ	CRENBL
PBE0	5.13/0.64	5.17/0.92
B3LYP	4.98/0.65	5.13/0.91
PBE	4.86/0.92	5.03/1.20

Table 5.2: Basis set and functional effects on $Ir(ppy)_3$ IP/EA values. All numbers are in eV.

not affect the hole transfer from the host to the guest, because the guest $-IP$ (-5.10 eV) is much higher than the host $-IP$ in spite of the broader distribution. Therefore, it is always a thermodynamically downhill process for the hole transfer to the guest.

From our QM/MMpol simulations, the averaged IPs/EAs of $Ir(ppy)_3$ and BTDF are 5.10/0.66 eV and 6.60/0.47 eV. The experimental values of the IP and EA are 5.1-5.3 eV and 1.9 eV for $Ir(ppy)_3$ [334, 335], while for BTDF the IP and EA are 6.54 eV and 1.68 eV [311]. Our calculated IPs agree very well with the experiments, but the EAs are underestimated. Possible reasons may be that the basis sets used in DFT are not large enough or the drude particle parameters in QM/MM simulations are not accurate. It should also be considered that the experiments have errors, especially for the measurement of EAs.

To understand the basis set and functional effects on IP/EA data, we did a small batch of QM/MM calculations employing larger basis sets and different functionals. All 15 guests and randomly selected 15 hosts in one snapshot are chosen as the test set. Tested basis sets include 6-31G*, 6-31+G* and 6-311G* for BTDF, as well as LANL2DZ and CRENBL [336] (larger than LANL2DZ) for $Ir(ppy)_3$. Tested functionals are commonly-used GGA functional PBE [43], and hybrid functionals PBE0 and B3LYP [49]. The averaged QM/MM results for the host and guest are shown in Tables 5.1-5.2. Larger basis sets (especially 6-31+G* and CRENBL) give EA values that are around 0.3 eV closer to the experiment compared to 6-31G* and LANL2DZ. Some particular functional (PBE) also gives better transport gap compared to experimental values. However, the IP/EA values of BTDF and $Ir(ppy)_3$ change in the same direction. The relative energy level relationship between the host and guest is not altered, so our conclusion about the charge recombination will not be affected. Thus, our basis sets and functional are sufficient for obtaining useful insights.

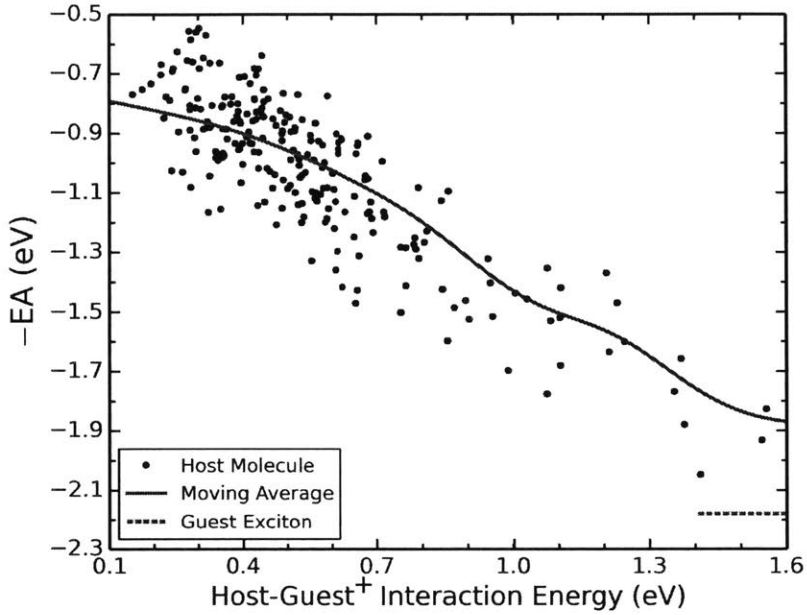


Figure 5-5: Electron affinities of the host near a guest cation as a function of the host-guest$^+$ interaction energy $|\Delta EA|$.

Since the driving force for the electron transfer step is uphill for many host-guest pairs, such a process cannot contribute significantly to efficient charge recombination on the guest. In this work, we would like to search for a recombination pathway where the reaction steps are all downhill. Transferring holes from host to guest first is the only candidate to achieve high recombination efficiency based on Figures 5-3 and 5-4. Therefore, it is reasonable to further consider the electron transfer from the host to the positively-charged guest. To obtain the host electron affinity near a guest cation, we perform QM/MMpol calculations on all host molecules in one snapshot where one guest is positively charged, shown in Figure 5-5. Note that the QM region is still one host molecule while the environment has changed due to the existence of the guest cation. Instead of the host-guest interaction energy, we use the host-guest$^+$ interaction energy here as the reaction coordinate. The only difference is that EA(on) in Eq. 5.4 is now the host EA from the calculation with the existence of a guest cation. Because the electron transfer to a guest cation would induce the formation of an exciton, it is necessary to compute the guest exciton energy level to compare with the host EA. We carry out QM/MMpol calculations on the guest utilizing the ROKS approach to obtain S_1 energy (E_{S_1}) and ($-IP + E_{S_1}$) is used as an estimation to the guest exciton

energy level. As there are fewer data points, we choose $\alpha = 30$ for the KDE procedure here (Eqs. 5.5-5.7). As seen from Figure 5-5, there is a strong correlation between the host EA and the host-guest$^+$ interaction energy: $-EA$ becomes significantly lower when the interaction energy is larger. It is also observed that all host electron affinities are above the guest exciton energy level (-2.18 eV), even for the ones that are very close to the guest. The electron transfer from the host to the guest cation becomes thermodynamically downhill due to the strong attraction of the guest holes. Therefore, this process completes the recombination pathway, and this is the only pathway that could possibly contribute to efficient electron-hole recombination directly on the guest.

5.2.3 Charge recombination pathways

To understand the distribution in Figures 5-3 and 5-4, we relate the host-guest interaction to the host-guest distance and their relative orientations. In a QM/MMpol calculation of E_+ and E_- of a host molecule, the dominant electrostatic interaction is charge-dipole interaction between the host and its surroundings, while for E_0 the dipole-dipole interaction dominates. Meanwhile, there are many possible relative molecular orientations between the host and guest, which results in a distribution of host EA and IP. When the host-guest interaction energy for a host molecule is large, at least one guest is close to the host. As a result, the molecular environment of this host is different from its bulk phase, as the guest has a distinct dipole moment from the host. In our model, $Ir(ppy)_3$ has a larger dipole (6.18D) than BTDF (0.87D) based on the DFT calculations. Therefore, the host near the guest has a broader distribution of EA and IP than in the bulk phase, as shown in Figures 5-3 and 5-4.

In Figure 5-5, a guest cation exists in the environment. Thus, the dominant electrostatic interaction between the host and guest becomes charge-charge interaction for E_- and dipole-charge interaction for E_0. Correspondingly, the EA distribution has a very strong correlation with the host-guest$^+$ interaction energy because the electrostatic interaction is much stronger than in Figures 5-3 and 5-4.

To better understand how charges recombine on the guest, we present Figure 5-6 and Figure 5-7 to explicitly describe the related processes. In Figure 5-6, we demonstrate the charge migration from the host to the guest through the direct charge trapping, while the electron transfer to the guest cation is shown in Figure 5-7. As shown in Figure 5-6, the

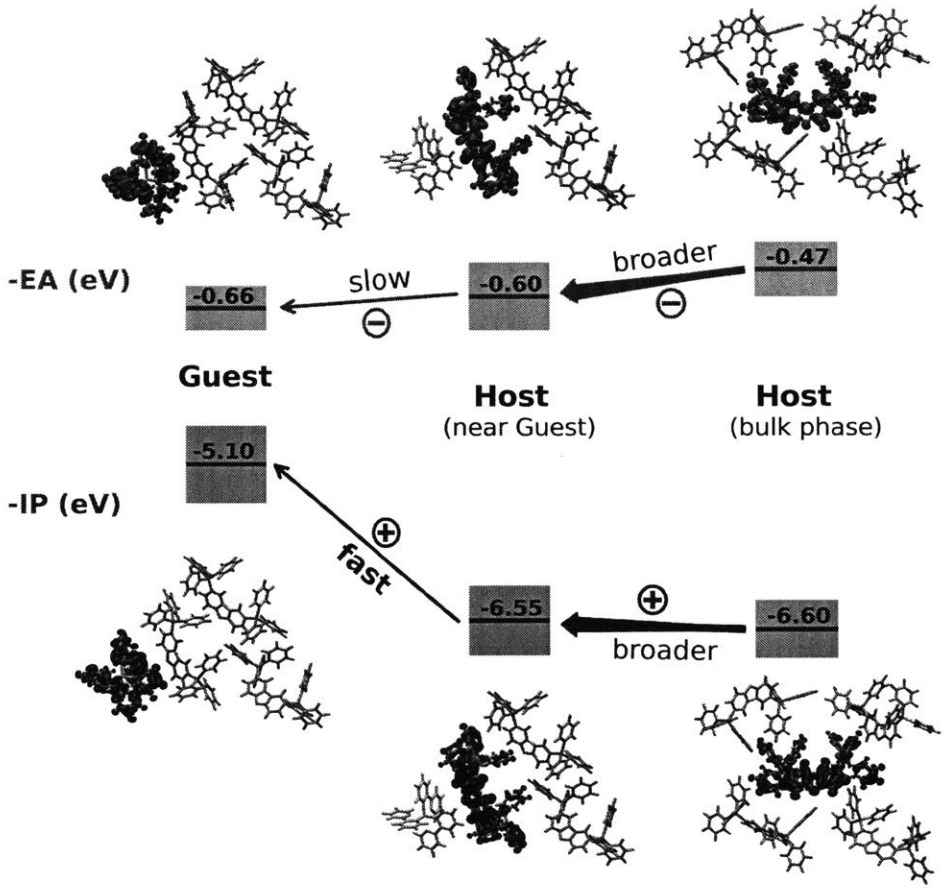


Figure 5-6: Charge recombination on the guest through direct electron and hole trapping. The standard deviations are shown by the blue/red regions around the energy levels.

charge migration to the guest can be divided into two steps: (1) from the bulk phase host to the host close to the guest; (2) from the nearby host to the guest. As the average EA and IP is similar between the host in the bulk phase (small host-guest interaction) and the host near the guest (large host-guest interaction), the first step is just as charge migration in the host material. However, the second step is different for the electron and hole transfer. For the electron transfer, as the host becomes closer to the guest, the host $-EA$ is slightly lower and its distribution becomes broader. As a result, many electrons get trapped on the host whose $-EA$ is below that of the nearby guest. Therefore, the electron migration is not as fast as from the host bulk phase to the guest. On the other hand, the hole transfer remains fast as the host is approaching the guest, because it is impossible to trap holes on the host whose $-IP$ is always much lower than that of the nearby guest.

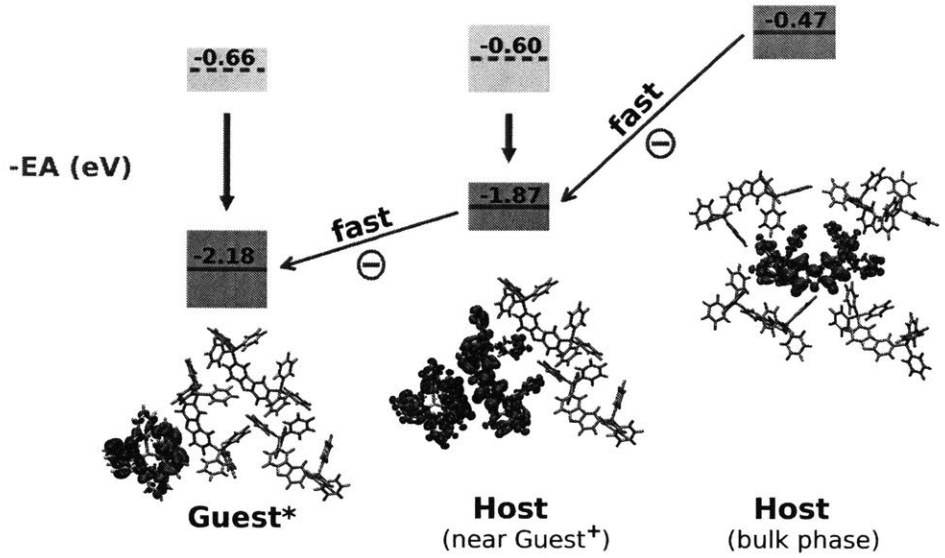


Figure 5-7: Charge recombination on the guest through electron transfer to the guest cation. The standard deviations are shown by the red regions around the energy levels.

It is clear that the hole transfer is the only downhill process to transfer charge from host to the neutral guest, we therefore consider the electron transfer from host to the guest cation. In Figure 5-7, owing to the existence of the hole on the guest, the electron energies of the guest and the nearby host both become lower. Furthermore, the guest $-EA$ lowers more than the host, causing the guest electron levels to be below the host electron levels. As a result, the electrons would not be trapped by the host and can recombine with the holes on the guest quickly.

For this specific host-guest system, we find one possible charge recombination mechanism to be the direct charge trapping. This finding could be significant for designing efficient host-guest systems. If this mechanism is dominant in a host-guest system, fast and well-balanced charge transfers should be the topmost consideration in experiments. Taking our system as an example, the electron energy of the host should be high enough in order to make the direct electron transfer as fast as the hole transfer. In addition, although the large offset between host and guest IPs gives fast hole transfer, it may create a narrow recombination zone in the emission layer which is far away from the electron transport side. As a result, this imbalance of hole and electron transfer would cause the degradation of OLED devices [337]. Thus, to make the offset of energy levels appropriate is also helpful. However, we also note that this mechanism may not be universal in all host-guest systems. One needs to do

analysis on more host-guest combinations to gain a comprehensive understanding.

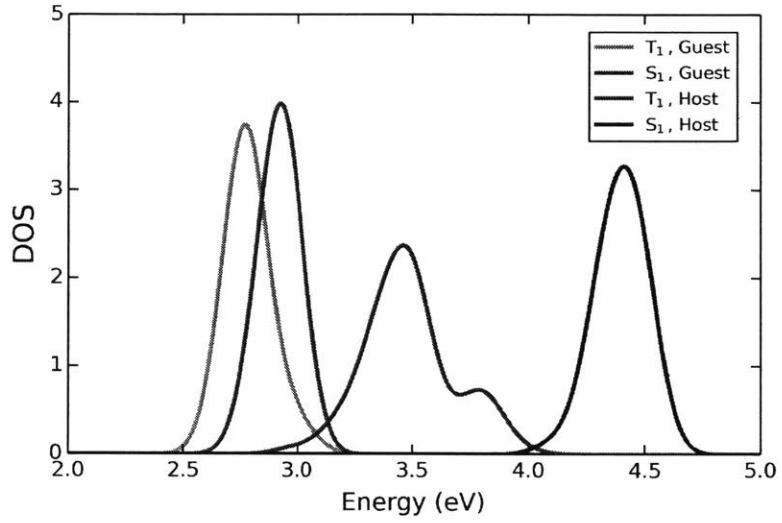
5.3 Mechanism of host exciton quenching

We have now studied the charge recombination mechanism in this BTDF-Ir(ppy)$_3$ PhOLED and found a charge trapping route existing in this system for direct exciton formation on the guest. However, it is still not clear how host excitons are quenched in this system. We herein examine all possible energy and charge transfer routes from excited host to the guest to study how host excitons can be efficiently harvested for guest emission. Using the MD simulation trajectory, we perform polarizable QM/MM (QM/MMpol) calculations on the ground and excited states of the host (75 molecules in 2 snapshots) and guest (15 molecules in 10 snapshots).

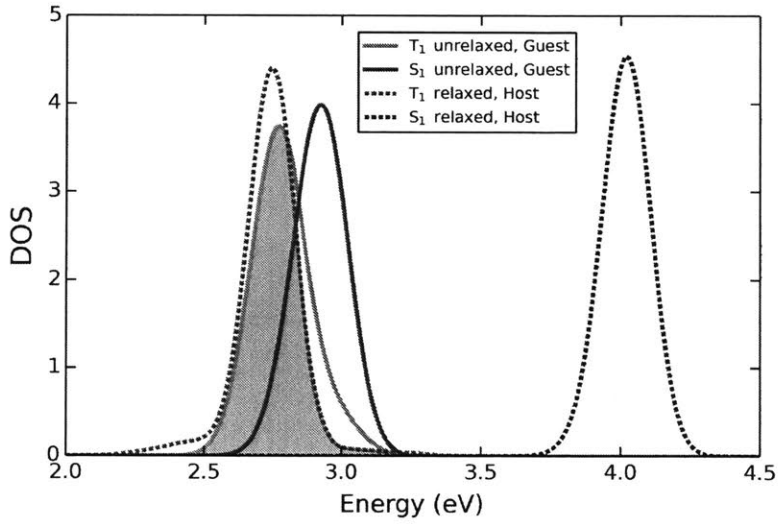
5.3.1 Energy transfer from excited host to guest

The lowest excited-state energies at MD-sampled ground-state geometries are calculated first (Figure 5-8a), to which we will refer as unrelaxed exciton energies. Unrestricted DFT (UDFT) and a restricted open-shell Kohn-Sham (ROKS) method [59] are employed to compute the T_1 and S_1 states respectively. This scheme has been shown to provide accurate prediction of excited-state energies for OLED molecules [338]. These QM/MMpol excited-state energies (150 data points for each state) are then used to construct probability distributions by applying a kernel density estimation (KDE). Oscillator strengths are not considered in Figure 5-8 as T_1 states are dark, making this figure a density of states (DOS) plot. As seen in Figure 5-8a, the unrelaxed T_1 and S_1 energies of BTDF are higher than those of Ir(ppy)$_3$ as expected, which guarantees that guest excitons would not transfer energy back to the host.

To investigate the energy transfer mechanism from excited host to the guest, we then simulate T_1 and S_1 energies of BTDF at its optimized excited-state geometries (Figure 5-8b), to which we will refer as relaxed exciton energies. Note that the unrelaxed/relaxed S_1 exciton energy is equivalent to the S_1 vertical absorption/emission energy. As MD simulation only samples ground-state structures, the excited-state geometry optimization is further performed for the host within the non-polarizable QM/MM (QM/MMnonpol) framework where drude charges are turned off. To reduce the computational cost, we freeze geometries



(a)



(b)

Figure 5-8: (a) Simulated DOS distributions for unrelaxed T_1 and S_1 excitons on the host and guest. (b) The DOS distributions for relaxed excitons on the host compared to the DOS distributions for unrelaxed excitons on the guest.

of all MM molecules and only optimize the geometry of the selected host molecule in the T_1 or S_1 state. QM/MMpol calculations are then carried out on the optimized excited-state geometries to obtain relaxed exciton energies of the host.

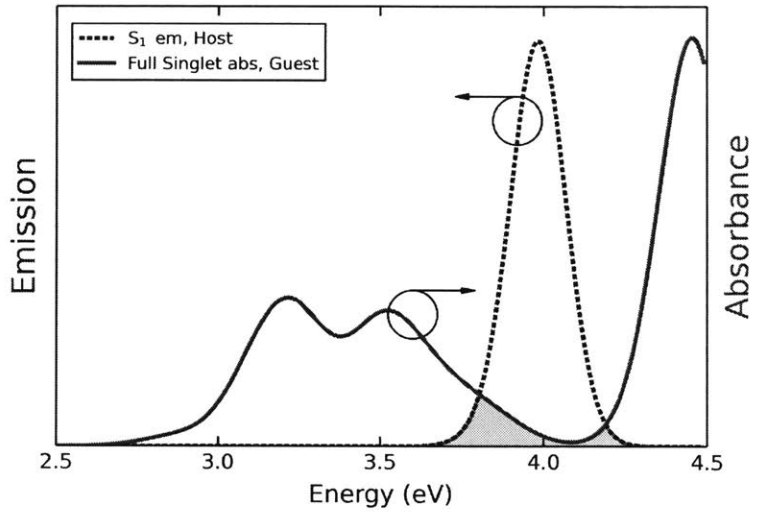
Since 75% of excitons generated from electron-hole recombinations are triplet [279], it

is important to harvest triplet host excitons to achieve high electroluminescence efficiency. Thus we discuss the triplet energy transfer first. In particular, only Dexter energy transfer (DET) is possible here, whose rate can be expressed as [299]:

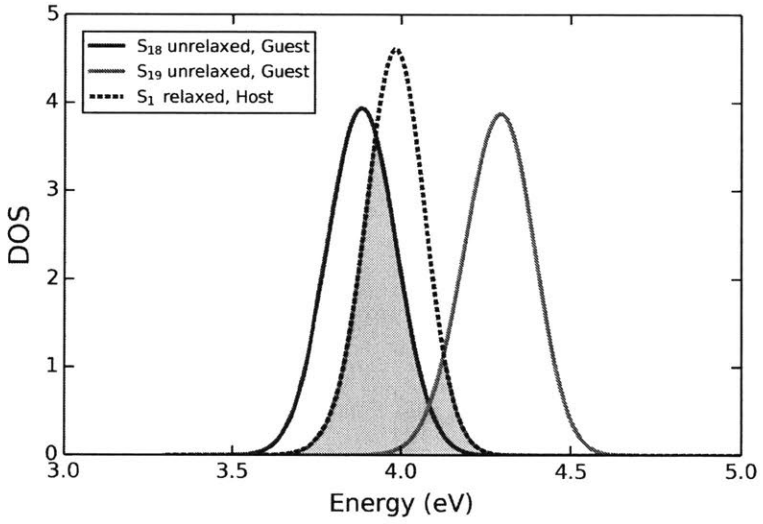
$$k_{DET} \propto \bar{J} \exp(-2R_{DA}/L), \tag{5.8}$$

where R_{DA} is the distance between donor (D) and acceptor (A), L is the sum of van der Waals radius. $\bar{J}$ is the normalized spectral overlap that is proportional to the overlap between DOS of donating and accepting states. It is demonstrated in Figure 5-8b that the DOS of relaxed T_1 excitons of BTDF overlaps substantially with the DOS of unrelaxed T_1 excitons of $Ir(ppy)_3$. This clearly suggests that triplet host excitons are allowed to transfer the excitation energy efficiently to the guest because the DET rate is proportional to the normalized spectral overlap $\bar{J}$. The DET rate is also affected significantly by the electronic coupling between donor and acceptor, indicated by its exponential relationship to R_{DA}/L. Nevertheless, when there is an appreciable ratio (for example, > 2 wt%) of guest molecules doped in host materials, most host molecules are adjacent to a guest molecule [306]. This wave function overlap term would not be the limiting factor then. This analysis therefore reveals that triplet host excitons in this system can lead to guest emission through efficient DET.

The quenching mechanism for singlet host excitons is more complicated. It is shown in Figure 5-8b that there is no DOS overlap between relaxed S_1 excitons of BTDF and unrelaxed S_1 excitons of $Ir(ppy)_3$, making the energy transfer between these two states impossible. Meanwhile, it is known that the host normally does not emit in PhOLED devices so singlet host excitons must be quenched. Although $Ir(ppy)_3$ S_1 absorption energy is substantially lower than BTDF S_1 emission energy, it is still possible for singlet host excitons to transfer energy to the guest if $Ir(ppy)_3$ has absorbing singlet states with energy levels around 4.0 eV. It is therefore interesting to take higher singlet excited states of the guest into consideration. To obtain the full absorption spectrum in the range of 2.5 to 4.5 eV, time-dependent DFT (TDDFT) is employed within the QM/MMnonpol framework to compute the first 25 singlet excited states. The full singlet absorption spectrum of guest and S_1 emission spectrum of host are shown in Figure 5-9a. As can be seen, our results are in good agreement with other TDDFT calculated and experimental spectra for $Ir(ppy)_3$ [339].



(a)



(b)

Figure 5-9: (a) Full singlet absorption spectrum of the guest (in the range of 2.5 to 4.5 eV) compared with S_1 emission spectrum of the host. (b) DOS distributions of unrelaxed S_{18} and S_{19} excitons on the guest and relaxed S_1 excitons on the host.

$Ir(ppy)_3$ has dense singlet energy levels in the shown spectrum range due to the existence of iridium and its symmetric structure, leading to a small overlap between the host S_1 emission spectrum and the guest singlet absorption spectrum. Particularly, Figure 5-9b indicates that the DOS of unrelaxed S_{18} excitons on the guest overlaps considerably with that of relaxed

host S_1 excitons.

As a dipole-dipole interaction process [340], the FRET rate can be expressed as [298]:

$$k_{FRET} \propto \frac{\kappa^2 J}{{R_{DA}}^6}, \tag{5.9}$$

where J is the spectral overlap and κ^2 is the dipole orientation factor. According to Eq. 5.9, FRET not only relies on the overlap of DOS distributions, but also requires strong absorption and emission intensities. However, the absorption intensity of $Ir(ppy)_3$ is weak around 4.0 eV (Figure 5-9a), where the peak of host S_1 emission spectrum is located on. FRET would thus not be effective due to the resulting small spectral overlap. On the other hand, DET is an electron exchange process whose rate only depends on the overlap of DOS distributions. Therefore, effective DET is allowed to occur from singlet host excitons to the high-lying singlet excited states of the guest, as suggested by the significant DOS overlap in Figure 5-9b. To conclude, in this case, it is DET rather than FRET that allows energy transfer from singlet host excitons to the guest.

5.3.2 Charge transfer from excited host to guest

In addition to the energy transfer mechanism, the charge transfer mechanism should also be explored, where host excitons break up and then reform on nearby guests. The key step is that a local host exciton is dissociated into an electron and a hole on a host-guest (H-G) pair, after which they can recombine on the guest. It is challenging to estimate the reaction energy of this process because the exciton binding energy and the energy difference between singlet and triplet excitons need to be evaluated. Thus, we propose to exploit thermodynamic cycles to address this problem and we focus on the exciton dissociation process, as shown in Figure 5-10. In Figure 5-10a, the first possible exciton dissociation process is demonstrated: it breaks up to form a guest cation and a host anion. Here, the binding energy between the guest cation and host anion may be large enough to stabilize the products and reduce the energy barrier considerably. It is therefore natural to investigate the reaction from $G\cdots H^*$ complex to $G^+\cdots H^-$ complex. Five thermodynamic steps are proposed to compute its reaction energy:

1. $G\cdots H^* \rightarrow G+H^*$: $G\cdots H^*$ complex is dissociated into independent G and H^*. $BE(G\cdots H^*)$ refers to the binding energy of G and H^*.

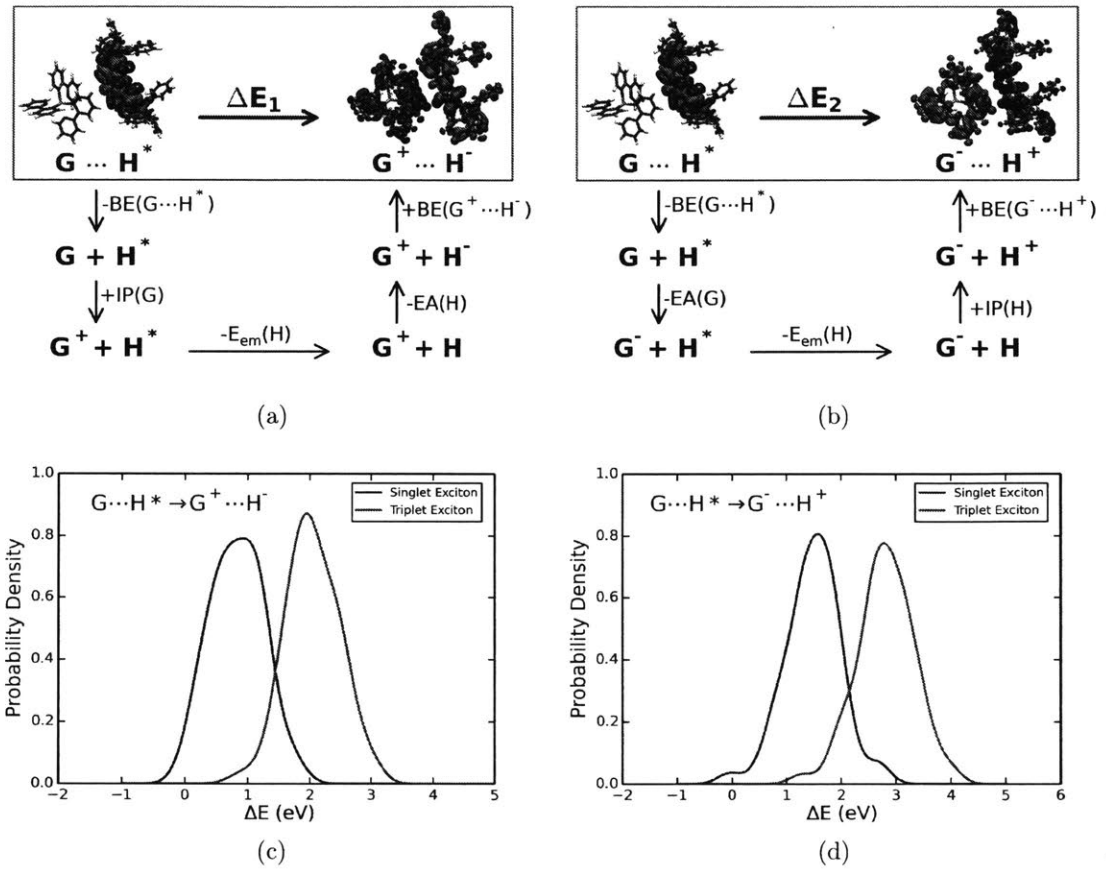


Figure 5-10: (a) Thermodynamic cycle for computation of the process that a host exciton is dissociated to form a host anion and a guest cation. (b) Thermodynamic cycle for computation of the process that a host exciton is dissociated to form a host cation and a guest anion. (c) The calculated probability distribution of reaction energy for the thermodynamic cycle in (a). (d) The calculated probability distribution of reaction energy for the thermodynamic cycle in (b).

2. $G+H^* \rightarrow G^+ + H^*$: a hole is added to the guest. IP(G) is ionization potential of G.

3. $G^+ + H^* \rightarrow G^+ + H$: the host exciton is quenched. Here singlet or triplet emission energy of H is calculated, and marked as $E_{em}(H)$.

4. $G^+ + H \rightarrow G^+ + H^-$: an electron is added to the host. EA(H) is electron affinity of H.

5. $G^+ + H^- \rightarrow G^+ \cdots H^-$: independent G^+ and H^- bind together to form $G^+ \cdots H^-$ complex. $BE(G^+ \cdots H^-)$ refers to the binding energy of G^+ and H^-.

As a result, the reaction energy of first exciton dissociation process can be expressed as:

$$\Delta E_1 = -\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*) + \mathrm{IP}(\mathrm{G}) - \mathrm{E_{em}}(\mathrm{H}) - \mathrm{EA}(\mathrm{H}) + \mathrm{BE}(\mathrm{G}^+\cdots\mathrm{H}^-). \tag{5.10}$$

Similarly, the reaction energy of the other exciton dissociation process (Figure 5-10b) where a guest anion and a host cation are generated can be calculated as:

$$\Delta E_2 = -\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*) - \mathrm{EA}(\mathrm{G}) - \mathrm{E_{em}}(\mathrm{H}) + \mathrm{IP}(\mathrm{H}) + \mathrm{BE}(\mathrm{G}^-\cdots\mathrm{H}^+). \tag{5.11}$$

To obtain EA, IP and $\mathrm{E_{em}}(\mathrm{H})$, we perform QM/MMpol calculations on 25 host-guest pairs in 2 snapshots. The distance between centers of mass of each host-guest pair is less than 12 Å to guarantee the electronic coupling for charge transfer is not a limiting factor. Besides, $\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*)$ and $\mathrm{BE}(\mathrm{G}^+\cdots\mathrm{H}^-)$ need to be calculated for the chosen host-guest pair. An efficient way to evaluate the binding energy for the host-guest pairs is to employ constrained DFT (CDFT) [65] with a polarizable continuum model (PCM) [341]. The binding energy $\mathrm{BE}(\mathrm{G}^+\cdots\mathrm{H}^-)$ can be calculated as:

$$\mathrm{BE}(\mathrm{G}^+\cdots\mathrm{H}^-) = \mathrm{E}(\mathrm{G}^+\cdots\mathrm{H}^-) - \mathrm{E}(\mathrm{G}^+) - \mathrm{E}(\mathrm{H}^-). \tag{5.12}$$

To obtain $\mathrm{E}(\mathrm{G}^+\cdots\mathrm{H}^-)$, we use CDFT to constrain a positive charge on the guest and a negative charge on the host for a host-guest complex. $\mathrm{E}(\mathrm{G}^+)$ and $\mathrm{E}(\mathrm{H}^-)$ are energies of independent guest cations and host anions that can be obtained by normal DFT calculations. As electrostatic interaction is dominant here, we apply the PCM model in all DFT calculations with dielectric constant of the host ($\epsilon = 2.82$) [309] to accurately simulate the electrostatic environment. The binding energy $\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*)$ is more difficult to evaluate. It is hard to constrain the singlet excitation locally on the host molecule of a host-guest complex because the excitation energy of the host is higher than the guest. However, unlike in $\mathrm{G}^+\cdots\mathrm{H}^-$, $\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*)$ only involves two neutral species, where H^* provides a similar electrostatic environment as H. Therefore, we estimate $\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*)$ using the binding energy between ground-state G and H ($\mathrm{BE}(\mathrm{G}\cdots\mathrm{H})$):

$$\mathrm{BE}(\mathrm{G}\cdots\mathrm{H}^*) \approx \mathrm{BE}(\mathrm{G}\cdots\mathrm{H}) = \mathrm{E}(\mathrm{G}\cdots\mathrm{H}) - \mathrm{E}(\mathrm{G}) - \mathrm{E}(\mathrm{H}), \tag{5.13}$$

where E(G$\cdots$H), E(G) and E(H) are all computed using normal DFT/PCM.

The results are presented in Figure 5-10c and 5-10d. As suggested by probability distributions, the reaction energies for both exciton dissociation processes are appreciably uphill by 0.8 eV or more. The triplet host exciton dissociation process has higher energy barrier due to its lower energy. It is also noticed that the energy barrier is lower in the first dissociation process ($G\cdots H^{*} \rightarrow G^{+}\cdots H^{-}$) than in the second. However, even for singlet host excitons in first dissociation process, the energy barrier is still substantial, which is due to the large local exciton binding energy of the host. Therefore, we conclude this charge transfer mechanism for quenching host excitons is unlikely to be efficient. Because exciton binding energies are similar in OLED molecules, this mechanism is probably inefficient for most OLED host-guest systems unless the band offset is massive.

5.4 Conclusions

In this chapter, we investigate the charge recombination and energy transfer in a model PhOLED host-guest system consisting of $Ir(ppy)_3$ and BTDF employing the QM/MM technique. Using IPs and EAs of the guest and host from QM/MM calculations, we introduce a new reaction coordinate to account for the influence of the guest on the energetics of the host, which is also correlated with the distance between them. We show that the larger host-guest interaction results in broader distributions of the host IP and EA. Then we describe a molecular picture of charge recombination on the guest through a charge trapping route: holes are transferred to the guest first, after which the electrons are attracted by the holes to form excitons on the guest. In addition, we reveal that energy transfer is the primary quenching mechanism for host excitons, while the charge transfer mechanism is unlikely to occur because of high energy barrier for exciton dissociation. Triplet host excitons are harvested through effective Dexter energy transfer to the T_1 state of guest due to excellent spectral overlap. Singlet host excitons, on the other hand, are forbidden from transferring energy to the S_1 state of the guest. Through more detailed analysis, we demonstrate that some of high-lying singlet states of guest overlap substantially with host S_1 state and therefore play a significant role in energy transfer for utilizing singlet host excitons.

We note that mechanisms uncovered here may not be general to all OLED host-guest systems. For example, the IP/EA energy level alignment in different host-guest system may

be different, which can lead to other charge recombination pathways. Meanwhile, the DET mechanism for quenching triplet host excitons may not be universal as the relaxation of triplet states will largely affect the spectral overlap between the host and the guest. The singlet exciton energy transfer will depend strongly on the presence and absorption intensity of high-lying excited states of the guest.

Nevertheless, this study provides important insights into better host-guest design. The charge recombination results suggest that consideration of the energy level change in a host-guest mixture environment may be needed to gain more accurate insight into the energy level alignment. Fast well-balanced charge trapping routes for host-guest systems may be critical for improving the device efficiency. Our host exciton quenching study suggests efficient energy transfer from host to guest can be utilized as an alternative strategy to the energy level alignment. In such case, charge trapping in the guest is not necessary so that better charge carrier mobility and lower driving voltage can be achieved [305]. We also highlight the necessity to evaluate high-lying excited states of the guest in host-guest design because they may be crucial in facilitating energy transfer in host-guest systems. Our future work should study more different host-guest combinations, take host-guest excimers into account, and investigate the effects of intersystem crossing and non-radiative decay on the exciton quenching mechanisms.

Using the condensed phase simulation of host-guest OLED as an example, we illustrate how to properly utilize DFT methods to model electronic properties of complex chemical systems. The study on the charge recombination mechanism in OLEDs emphasizes the importance of incorporating environment effects into DFT simulations, which is crucial for capturing the energy level changes in different molecular environment. In addition, the study on the host exciton quenching mechanisms points out the significance of utilizing an accurate and efficient excited-state DFT method. In conclusion, existing DFT methods provide a powerful tool for simulating complex chemical systems, but better DFT-based methods for condensed phase simulation are still needed. We will further discuss the future directions in this aspect in Chapter 7.

5.5 Acknowledgements

We thank Dr. Denis Andrienko for providing the BTDF force field. We thank Dr. Shuhao Wen, Dr. David MaMahon and Dr. Piotr de Silva for helpful discussions.

Chapter 6

Computational design of thermally activated delayed fluorescence materials

6.1 Introduction

Since the first report by Tang and VanSlyke in 1987 [27], multilayered organic light emitting diodes (OLEDs) have attracted interest for utilization in high efficiency illumination and flexible displays [273, 274, 276, 277]. OLEDs using fluorescent materials are limited to have low internal quantum efficiencies (IQEs) of 25% due in part to the inherent limitation of electrical excitation, which generates singlets and triplets in a 1:3 ratio [278, 279, 280]. High quantum yield OLEDs with Ir or Pt phosphorescent materials have been intensely investigated for the last several decades and now achieve 100% IQE [281, 282, 283, 284]. Although phosphorescent materials have defined the present state of OLED technology, there are significant issues including cost, stability of blue emitters, and strong triplet-triplet annihilation at high current density [342]. As a result of recent efficiency increases, thermally activated delayed fluorescence (TADF) has become a viable alternative for harvesting both singlet and triplet state in OLEDs [28, 343, 344]. TADF is based on reversible intersystem crossing from thermally equilibrated triplet and singlet excited states, and competitive luminescence from the singlet states, as shown in Figure 6-1. If nonradiative pathways are negligible then TADF can achieve 100% electroluminescence IQE [290]. An advantage of TADF materials

is that they can be purely organic materials and do not require expensive metallic elements such as Pt and Ir, which also offers new design opportunities for both molecules and devices.

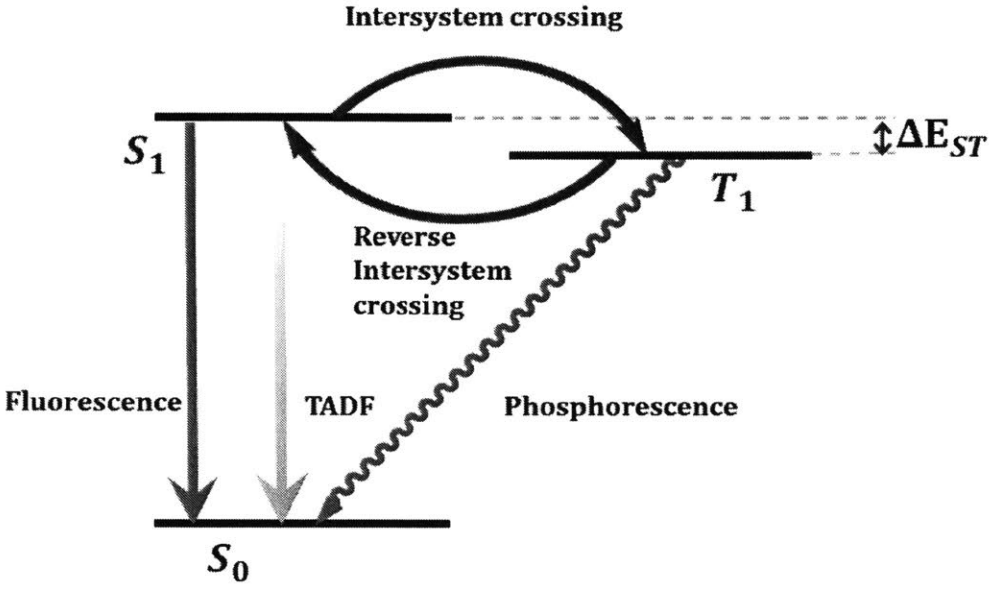


Figure 6-1: Illustration of thermally activated delayed fluorescence.

TADF is only possible at appreciable rates if the energy gap (ΔE_{ST}) between the S_1 and T_1 states is smaller than or comparable to $k_B T$. This can usually achieved by the spatial separation of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), which minimizes the exchange interaction between the two singly occupied orbitals in excited states and thus leads to charge transfer (CT) excited states. Meanwhile, efficient radiative decay from the S_1 state to the ground state S_0 is also required, which is usually manifested by large transition dipole moment or oscillator strength (f). Although for a compound to possess a large f value requires substantial overlap between HOMO and LUMO, a small ΔE_{ST} can be realized only when this overlap is diminished. Therefore, substantial efforts have been made to utilized molecular design strategies to solve this dilemma [345], which are often based on modifications of donor-acceptor (D-A) type fluorophores [346, 344, 347].

In addition to the ΔE_{ST}-f dilemma, OLED applications also frequently require emitted radiation of a particular color, thereby constraining suitable values of emission energy (E_{emit}) to a narrow interval. It is therefore useful to have computational tools for prediction of photophysical properties like ΔE_{ST}, E_{emit}, quantum yields, etc., of molecules with CT

excited states, as it allows screening of molecules for use in TADF-based OLEDs. Unfortunately, many of the molecules of interest are too large (>100 atoms) to be studied with high-level *ab initio* wavefunction-based methods like CASPT2 [61] or coupled-cluster [7], which are known to give quite accurate results for small organic molecules [62]. As a result, density functional theory (DFT)-based approaches are the only viable post-Hartree-Fock computational tools for studying such molecules.

In this chapter, we describe the application of density functional theory on the computational design of TADF materials. Instead of complicated condensed phase simulations, we focus on predicting excited-state state properties of single TADF molecule in vacuum, which is also challenging for common excited-state DFT methods. By showing several successful computational designs of TADF materials, we illustrate the usefulness of DFT simulations in predicting photophysical properties for experiments. Meanwhile, we note that conventional excited-state DFT methods still need improvements, leading to our development of a new computational protocol for computing TADF properties based on a restricted open-shell Kohn-Sham (ROKS) approach.

6.2 Computational screening of TADF materials

In this section, we describe our efforts to design new TADF materials, in collaboration with experimentalists. To achieve high TADF efficiency, we adopt two different molecular design strategies (Figure 6-2): (1) through-space interaction based on donor-acceptor homoconjugation [348]; (2) dihedral angel tuning of D-A type fluorophores [349].

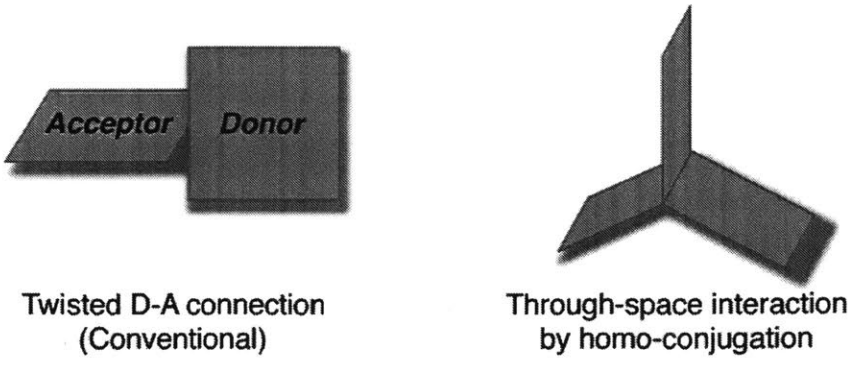


Figure 6-2: TADF design strategies explored in this work.

6.2.1 Computational details

The most standard DFT-based excited state method is linear response time-dependent density functional theory (TDDFT) [63, 64]. One of the earliest protocols for predicting TADF properties was designed by Adachi and co-workers [350, 351], who attempted to obtain estimates for ΔE_{ST} from the difference between S_1 and T_1 energies obtained at S_0-optimized geometries with TDDFT employing semi-local hybrid functionals like B3LYP [49] or PBE0 [328]. This approach, however, may not be effective for molecules with very high CT character, as semi-local TDDFT systematically underestimates the energy of CT states in general [352, 353, 354, 355]. Nevertheless, because of its simplicity and efficiency, it can still serve as a computational screening tool for the TADF molecular design.

Therefore, in this work, we first use TDDFT calculations to guide our TADF materials design. All DFT computations are carried out using the Q-Chem 4.1 software package [230]. The gas-phase ground state geometry optimizations for all TADF molecules are performed using the B3LYP exchange-correlation functional in the 6-31G* basis set [323]. Then time-dependent DFT is adopted at the same level to estimate the excitation energies.

6.2.2 Through-space interaction from homoconjugation

Although the most commonly employed TADF design strategy is to twist the donor-acceptor connection [344, 356] (left panel in Figure 6-2), an alternative approach is a through-space interaction where electronic systems are in communication by homoconjugation [357] but are sufficiently separated to create a small singlet-triplet energy gap ΔE_{ST} (right panel in Figure 6-2). The design we report here places the donor and acceptor on the different fins of a triptycene scaffold. These structures display homoconjugation and many triptycene derivatives display intrinsically high thermal stability, which is critical to OLED manufacturing and operation [358, 359].

We design the donor-acceptor triptycenes, TPA-QNX(CN)2 and TPA-PRZ(CN)2, as novel TADF materials (Figure 6-3). The triphenylamine (TPA) functions as the donor and dicyanoquinoxaline (QNX) or dicyanopyrazine (PRZ) is the acceptor. We perform TDDFT calculations to estimate the singlet-triplet energy gap ΔE_{ST} and oscillator strength f for these two compounds and first show their molecular orbitals in Figure 6-4. Figures 6-4a and 6-4b reveal that the HOMO is located on the TPA moiety, while the LUMO is localized on

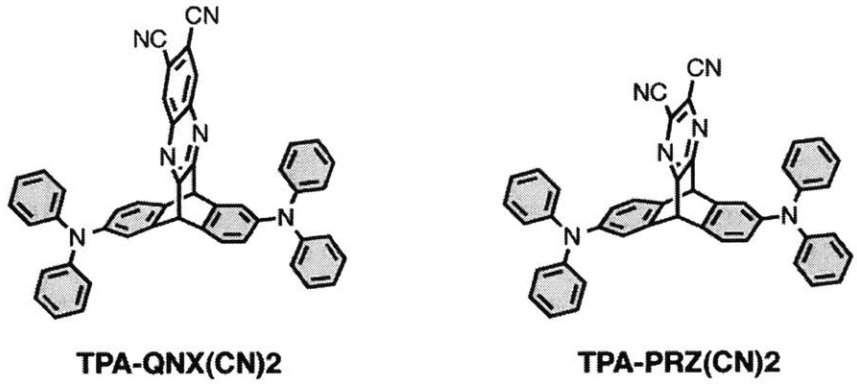


Figure 6-3: Chemical structures of triptycene-based TADF materials.

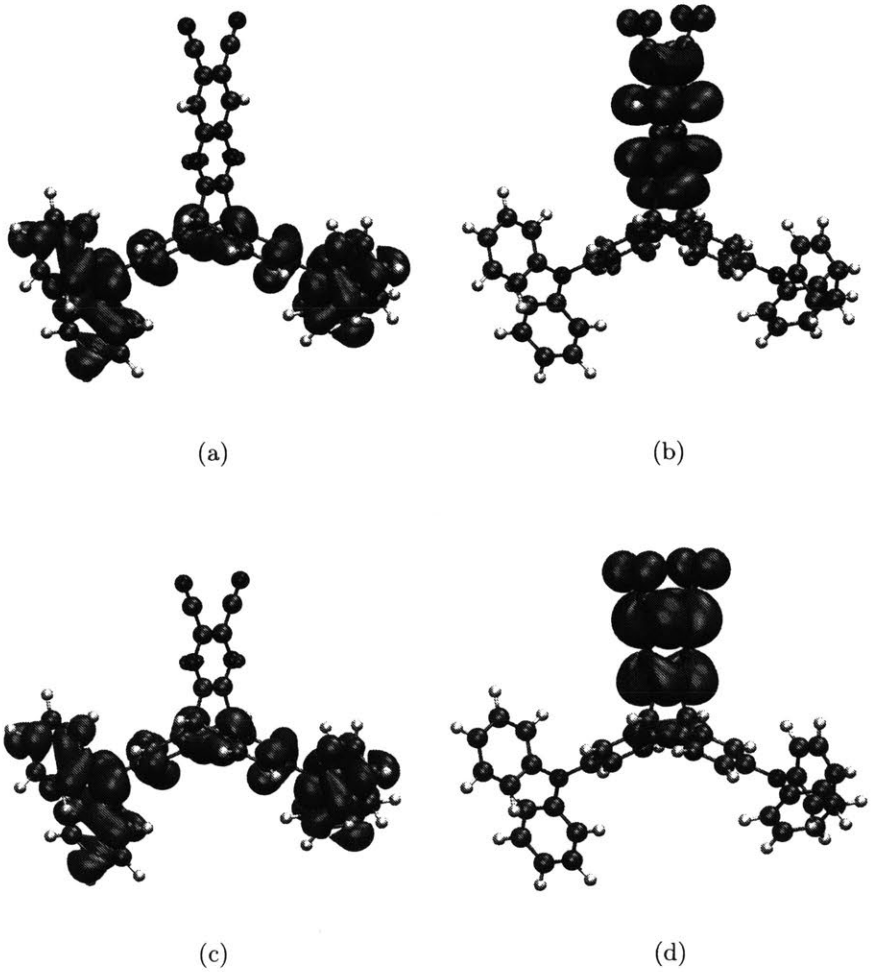


Figure 6-4: Molecular orbitals of the TADF compounds. (a) HOMO of TPA-QNX(CN)2. (b) LUMO of TPA-QNX(CN)2. (c) HOMO of TPA-PRZ(CN)2. (d) LUMO of TPA-PRZ(CN)2.

the QNX moiety, suggesting that the HOMO and LUMO are very well separated on TPA-QNX(CN)2 from our design. Similarly, a good HOMO-LUMO separation is also observed for TPA-PRZ(CN)2 in Figures 6-4c and 6-4d.

Compound	S_1 (eV)	T_1 (eV)	ΔE_{ST} (meV)	f
TPA-QNX(CN)2	2.25	2.14	111	0.084
TPA-PRZ(CN)2	2.33	2.25	75	0.008

Table 6.1: TDDFT calculation results of TPA-QNX(CN)2 and TPA-PRZ(CN)2.

We then present the TDDFT results for these two compounds in Table 6.1. As can be seen, the singlet-triplet energy gap ΔE_{ST} of TPA-QNX(CN)2 is only 111 meV, indicating that the small HOMO-LUMO overlap leads to small ΔE_{ST}. Meanwhile, the oscillator strength f of TPA-QNX(CN)2 is 0.084, which is substantial considering such a small HOMO-LUMO overlap. This is probably a result of homoconjugation interactions between the sp^2 C's attached to the bridgehead carbons. Thus, the small singlet-triplet gap and substantial oscillator strength suggest TPA-QNX(CN)2 is a good candidate as a TADF emitter. TPA-PRZ(CN)2 has a even smaller ΔE_{ST} of 75 meV, but the oscillator strength drops to only 0.008. This is expected to cause the singlet radiative decay rate to be too small and may yield worse TADF efficiency than TPA-QNX(CN)2 despite the smaller ΔE_{ST}.

Then we conduct experiments to test the photophysical and device properties of these two new TADF emitters. As the focus of this section is the computational design, we only present the TADF-OLED device experimental data. For more details the readers are referred to Ref. [348]. OLED devices using the newly-designed triptycenes as the emissive element are fabricated within the following architecture: ITO (132 nm)/MoO3 (5 nm)/TcTa (30 nm)/10 wt% triptycene:mCP (30 nm)/TmPyPb (40 nm)/LiF (0.8 nm)/Al (100 nm) (Figure 6-5a). The electroluminescence spectra of OLED devices based on TPA-QNX(CN)2 and TPA-PRZ(CN)2 are shown in Figure 6-5b. The electroluminescence emission maxima are 573 nm (CIE = 0.45, 0.54) for TPA-QNX(CN)2 and 542 nm (CIE = 0.43, 0.55) for TPA-PRZ(CN)2. We obtain external quantum efficiencies (EQE) in OLEDs based on TPA-QNX(CN)2 with values up to 9.4% (Figure 6-5c, red line). This exceeds the highest EQE values for OLED devices based on simple fluorescence materials of around 5%. We attribute the increased performance of TPA-QNX(CN)2 to its TADF properties. Our highest EQE of OLEDs using TPA-PRZ(CN)2 as an emitter are 4.0% (Figure 6-5c, blue line). The lower EQE compared to the TPA-QNX(CN)2 OLED is consistent with our computational predictions

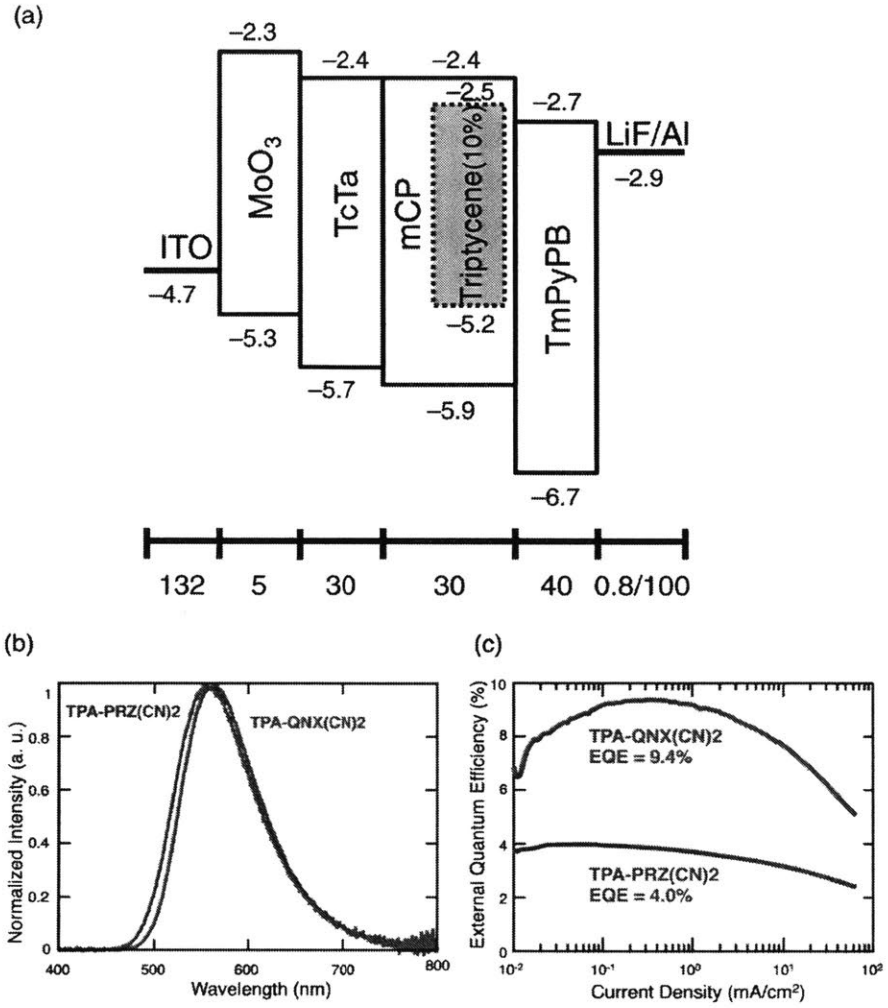


Figure 6-5: (a) Energy band diagrams and device structure of the OLED devices. The emissive layers are coevaporated with 10 wt% of triptycene molecules in host mCP. TcTa, Tris(4-carbazoyl-9-ylphenyl)amine; mCP, 1,3-Bis(N-carbazolyl)benzene; TmPyPb, 1,3,5-Tris(3-pyridyl-3-phenyl)benzene. The energy values are all given in eV. The thicknesses are in nm. (b) Electroluminescence spectra. (c) The EQE, current density of the OLED, using the triptycene emitters.

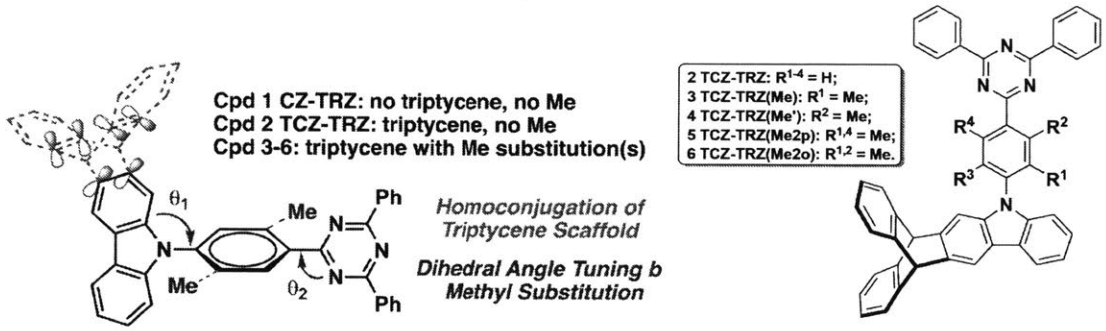
and we attribute the worse performance of the TPA-PRZ(CN)2 OLED to its low oscillator strength. This result suggests the importance of achieving balanced singlet-triplet gap and oscillator strength through molecular design.

In summary, we designed donor-acceptor triptycenes TPA-QNX(CN)2 and TPA-PRZ(CN)2 as novel TADF emitters. These newly designed triptycene-based TADF materials make use of the physical separation of the donor and acceptor groups on different fins of the triptycene scaffold. The intramolecular orbital overlap is accomplished by homoconjugation.

Multilayer OLED devices using these new triptycene emitters demonstrate yellow emission with high EQE up to 9.4%. The good performance of designed TADF materials suggests our TDDFT-based computational screening tool provides efficient and useful predictions of the TADF properties.

6.2.3 Dihedral angel tuning

In this section, we describe the successful transformation of a fluorophore CZ-TRZ into a series of deep blue TADF emitters by applying two design strategies: (1) homoconjugation through incorporation of a triptycene scaffold and (2) manipulation of dihedral angles by methyl substitution (Figure 6-6a). Although the parent molecule CZ-TRZ **1** emits deep blue light, most TADF molecules having this basic structure have substituents that shift the S_1 emission to sky-blue or green [344, 360, 361]. We surmise that homoconjugation could decrease ΔE_{ST} while maintaining a high S_1 energy and have shown that the triptycene framework can be utilized to form TADF emitters. Here, we hypothesize that the homoconjugation effect could be more generally applied if the triptycene scaffold could be incorporated into the carbazole donor as the triptycene-fused carbazole (TCZ). In addition to employing the TCZ moiety as the donor, we also carry out methyl substitutions at various positions of the phenyl linker in CZ-TRZ, leading to a series of new TCZ-TRZ compounds (Figure 6-6b).



(a) (b)

Figure 6-6: (a) Homoconjugation and dihedral angle tuning. (b) Chemical structures of Compounds **2-6**.

We perform time-dependent density function theory (TDDFT) calculations to estimate the singlet-triplet energy gap ΔE_{ST} and oscillator strength f of Compounds **1-6**, as shown

Compound	S_1 (eV)	T_1 (eV)	ΔE_{ST} (meV)	f	angle θ_1/θ_2 (deg)
1	3.05	2.70	346	0.315	51.5/0.9
2	2.95	2.66	292	0.276	51.6/0.8
3	2.88	2.76	116	0.081	69.5/0.4
4	2.99	2.72	271	0.243	51.5/25.1
5	2.94	2.83	108	0.073	69.6/26.9
6	2.99	2.90	89	0.053	69.1/34.1

Table 6.2: TDDFT calculation results of Compounds **1-6**.

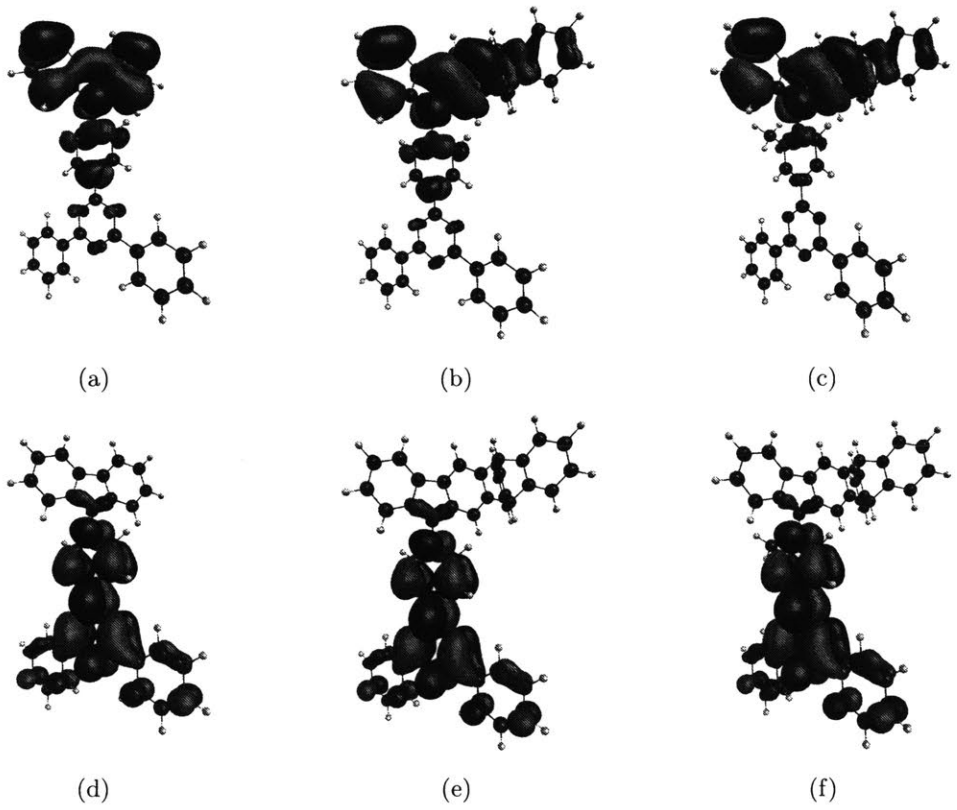


Figure 6-7: Molecular orbitals of Compounds **1-3**. (a) HOMO of Compound **1**. (b) HOMO of Compound **2**. (c) HOMO of Compound **3**. (d) LUMO of Compound **1**. (e) LUMO of Compound **2**. (f) LUMO of Compound **3**.

in Table 6.2. Compounds **1** and **2** are calculated to have similar values for f, implying that both should exhibit strong fluorescence. The estimated ΔE_{ST} of **2** (0.292 eV) is smaller than that of **1** (0.346 eV), while the S_1 energy of **2** remains high at 2.95 eV. Substituting hydrogen with methyl on the phenylene ring at a position *ortho* to the carbazole results in sterically induced twisting [356, 362]. As the dihedral angle θ_1, defined by the carbazole

plane and the phenylene plane (Figure 6-6a), changes from 51.6° in **2** to 69.5° in **3**, the ΔE_{ST} of **3** is further reduced to 0.116 eV. However, when the phenylene ring is substituted at a position *ortho* to the triazine as in **4**, the ΔE_{ST} stays almost unchanged at 0.271 eV despite an increase in dihedral angle θ_2, defined by the triazine plane and the phenylene plane (Figure 6-6a). As a result, the calculated values of f and ΔE_{ST} for **5** and **6** are similar to that of **3**. The HOMOs and LUMOs of compounds **1-3** are depicted in Figure 6-7. While all LUMOs localized on the triazine are qualitatively similar, the HOMOs of **2** and **3** extend to the neighboring phenyl rings because of the homoconjugation effect, which reduces HOMO and LUMO overlap at the phenylene compared to **1**. The overlap is further reduced in **3** because of the larger θ_1. In sum, our TDDFT calculations predict Compounds **3**, **5**, and **6** are likely to be efficient blue TADF emitters.

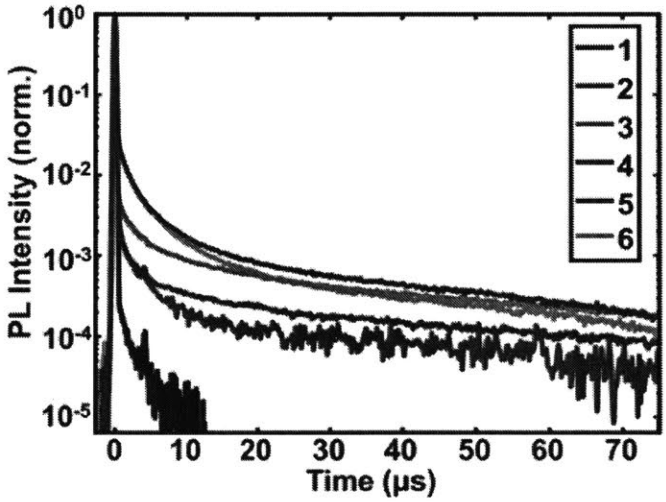


Figure 6-8: Transient decay of Compounds **1-6** (measured in codeposited film (15 wt%) in DPEPO).

We then carry out experiments to investigate the photoluminescence properties of the newly designed compounds in toluene and bis[2-(diphenylphosphino)phenyl] ether oxide (DPEPO) films (doped at 15 wt%). We first measure the PL quantum yields (PLQYs) of Compounds **1-6** to evaluate their potential performance in OLED devices. In solution, Compounds **2** (PLQY = 0.77) and **4** (0.80) exhibite high PLQY values comparable to those of **1** (0.78), while the values for Compounds **3** (0.60), **5** (0.46), and **6** (0.47) are much lower. This is consistent with the trend of TDDFT-calculated oscillator strength values in Table 6.2. Time resolved photoluminescence measurements are then performed on doped

films (Figure 6-8). For prompt fluorescence, Compounds **1-6** show similar decay lifetime ranging from 7.2 to 10.6 ns. While **1** barely has any delayed emission, Compounds **2-6** clearly exhibit delayed emission with lifetimes ranging from 37 to 58 μs. More importantly, the delayed component is more prominent in compounds with θ_1 (**3**, **5**, and **6**) larger than that of the others (**2** and **4**). This observation is again consistent with our TDDFT prediction: increasing dihedral angle θ_1 results in smaller ΔE_{ST} and thus leads to stronger TADF character.

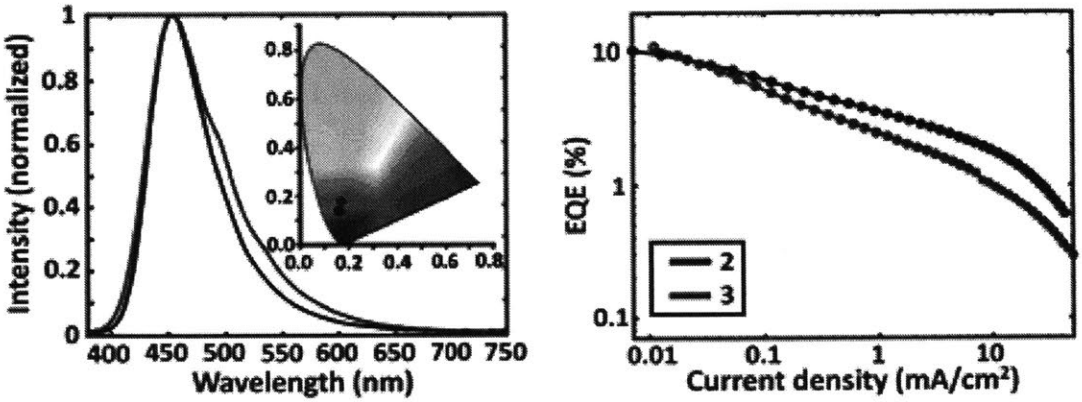


Figure 6-9: (a) EL spectra of **2** (black) and **3** (red) with their location in the CIE color space at 50 cd/m^2. (b) EQEs of **2** and **3** as a function of current density.

Finally we present the electroluminesence (EL) data of OLED devices employing Compounds **2** and **3** as emissive dopants in Figure 6-9. We obtain maximum external quantum efficiencies (EQEs) of 10.4 and 11.1% in devices employing **2** and **3**, respectively (Figure 6-9b). These values exceed the theoretical limit of typical fluorescence materials and are 2.5 times higher than the 4% EQE for **1** [363], providing additional evidence that these materials exhibit TADF. However, the devices using Compounds **4-6** as emitters showed lower EQEs, which is attributed to their weak absorption, low PLQYs, and rapid degradation under electrical excitation. The EQEs and CIE values for Compound 2 and 3 at 50 cd/m^2 are 3.4%, (0.159, 0.142) and 2.0%, (0.170, 0.179), respectively, suggesting their usefulness as deep blue OLED emitters. The readers are referred to Ref. [349] for more experiment details.

In summary, we successfully transform fluorescent emitter **1** to TADF emitters **2** and **3** with EQEs surpassing 10% while maintaining the deep blue emission. Our results indicate that the introduction of a homoconjugative triptycene scaffold enhanced the TADF proper-

ties by effectively reducing the overlap between HOMO and LUMO, while the dihedral angle tuning does not yield significant improvement of device performance because it has conflicting effects in PLQYs and ΔE_{ST}. Nevertheless, the good agreement between the theory and experiment data indicate the TDDFT-based computational method can be an efficient tool to estimate TADF properties for materials without too strong CT character ($\Delta E_{ST} > 100$ meV).

6.3 Prediction of TADF energetics based on a ROKS approach

We have shown that TDDFT calculations based on S_0 optimized geometries and typical hybrid functionals (e.g., B3LYP) provide efficient predictions of the singlet-triplet energy gap ΔE_{ST} and the S_1 emission energy for the triptycene-based and TCZ-TRZ-based TADF molecules. However, these molecules do not possess very strong CT character, which can still be treated by TDDFT with semi-local hybrid functionals at a reasonable accuracy. For molecules with very strong CT character, this TDDFT protocol tends to underestimate ΔE_{ST} [338]. Meanwhile, we also note that the accurate estimation of the S_1 emission energy is due to the cancellation of errors: TDDFT with B3LYP underestimates CT states [352, 353, 354, 355], while using S_1 absorption energy as the prediction for the S_1 emission energy results in overestimation. Therefore, more accurate computational protocols for efficient predictions of TADF energetics are still needed.

Another method proposed by Adachi and co-workers [364] tried to determine the ideal TDDFT functional for each structure by empirically calculating an "optimal" percentage of exact exchange from ground-state calculations that estimated the extent of CT. Any semi-local hybrid functional employing close to the calculated exact-exchange percentage was postulated to be effective for TDDFT studies. An alternative approach would be to use range-separated hybrid functionals [365, 366] that were developed for the purpose of performing TDDFT on CT states. This was also tested by Adachi and co-workers, who however reported that range-separated functionals like CAM-B3LYP [367] or LC-ωPBE [368] tended to overestimate absorption energies (E_{abs}) for common TADF molecules, indicating that the range separation parameters for such molecules were not optimal for the length scale of charge transfer in such systems. This was not particularly surprising, as such parameters are often strongly system dependent [369], although it is possible to "tune" them

for individual systems [370]. Recently, Penfold [371] and Bredas et al. [17] independently used tuned range-separated functionals to investigate TADF molecules and discovered that such an approach gave ΔE_{ST} values that are fairly consistent with experimental data.

There also exist time-independent excited-state DFT techniques like ΔSCF [372], which offer alternate routes for studying CT excited states. Such methods generally do not rely on linear response theory and can therefore be expected to not share the deficiencies of semi-local TDDFT with regard to CT states. In particular, it is possible to use a restricted open-shell Kohn-Sham (ROKS) approach [58, 59] to obtain energies of the S_1 state, which offers a new way to estimate emission wavelengths. Additionally, the Hohenberg-Kohn theorem [37] indicates that ground-state spin density functional theory (SDFT) should be able to estimate the energy of the first excited triplet (T_1) state, as it is the ground state within the subspace of all triplet electronic states. This indicates that a combination of ROKS and SDFT could also be employed to calculate ΔE_{ST}, thereby implying that such a combination could be useful as a fast and reliable computational screen for potential TADF molecules.

In this section, we address these questions by devising two computational protocols [338] that use a combination of ROKS and SDFT to estimate photophysical properties including E_{abs} (S_1 vertical absorption energy), E_{emit} (S_1 vertical emission energy), E_{0-0} (gap between the minimum S_1 energy and the minimum S_0 energy), and ΔE_{ST} (gap between the minimum S_1 energy and the minimum T_1 energy). Consequently, these protocols are compared with two TDDFT-derived protocols against a test set of 27 TADF compounds (Figure 6-10) with available experimental data [346, 364, 373, 374]. This reveals that while it is possible to use cancellation of errors to obtain some useful results from TDDFT, such approaches in general do not lead to accurate estimations of all desired parameters. On the other hand, semi-local hybrid functional-based ROKS/SDFT-based approaches lead to very accurate predictions of E_{emit}, ΔE_{ST}, etc., without having to resort to cancellation of errors or parameter optimization. Overall, it appears that ROKS with semi-local hybrid functionals provides a reliable and fast alternative to estimating properties of CT states and can thus be used as a screen for potential TADF molecules.

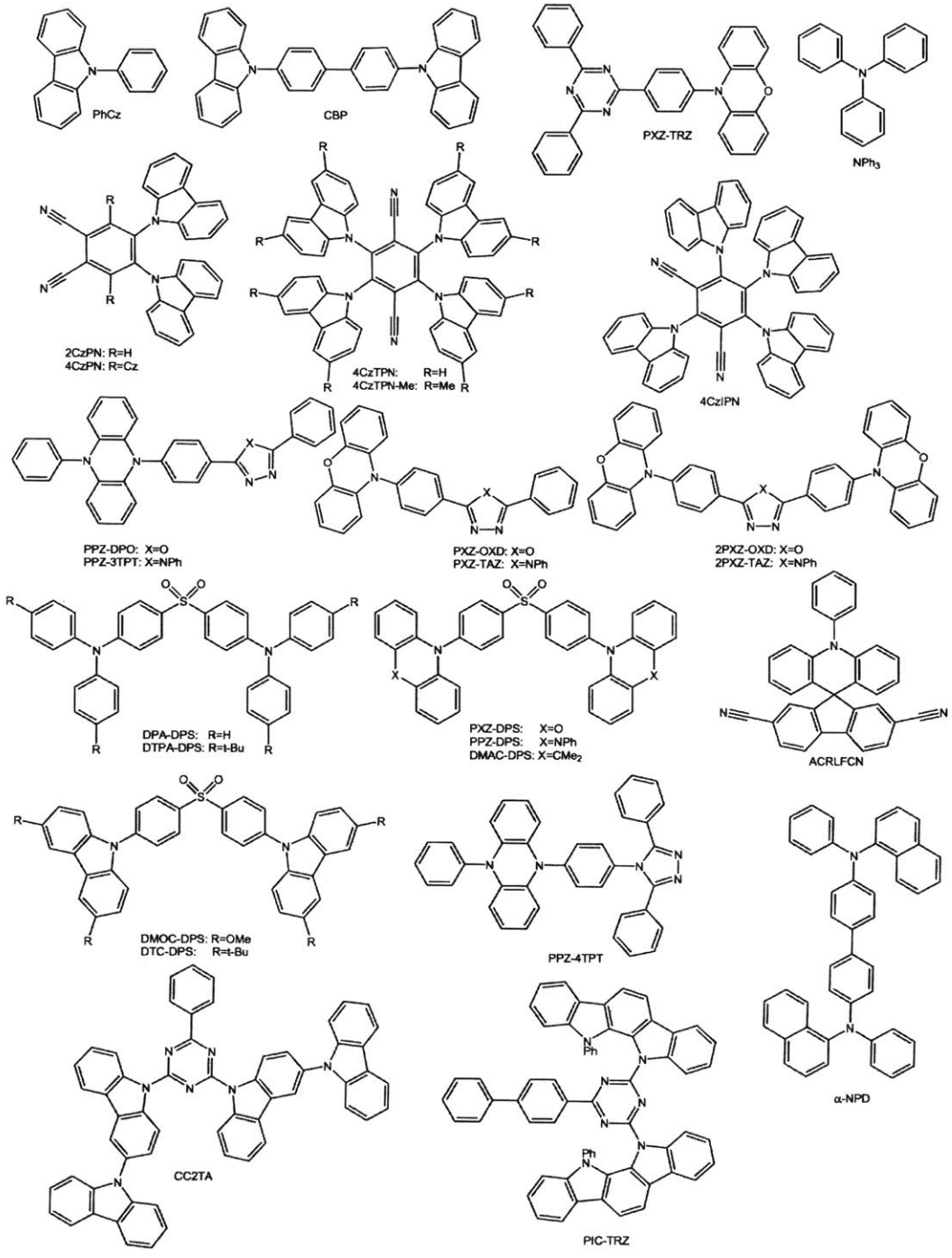


Figure 6-10: Structures of all the TADF molecules in the test set.

6.3.1 Computational protocols

All the calculations are done with the QChem 4.2 package [230], employing the PBE [43], B3LYP [49], PBE0 [328], and LC-ωPBE [368] (with the QChem default Coulomb attenuation parameter $\omega = 0.3$ bohr^{-1} and zero short-range HF exchange) functionals. Most calculations employed the 6-31G* basis set [323], although some B3LYP calculations are repeated with the larger cc-pVTZ basis [231] to investigate the basis set dependence of the parameters measured. The Tamm-Dancoff approximation (TDA) [375, 376, 377] is not invoked in TDDFT calculations. Solvent effects are not taken into consideration in this study nor is the effect of zero-point energy of molecular vibrations taken into account.

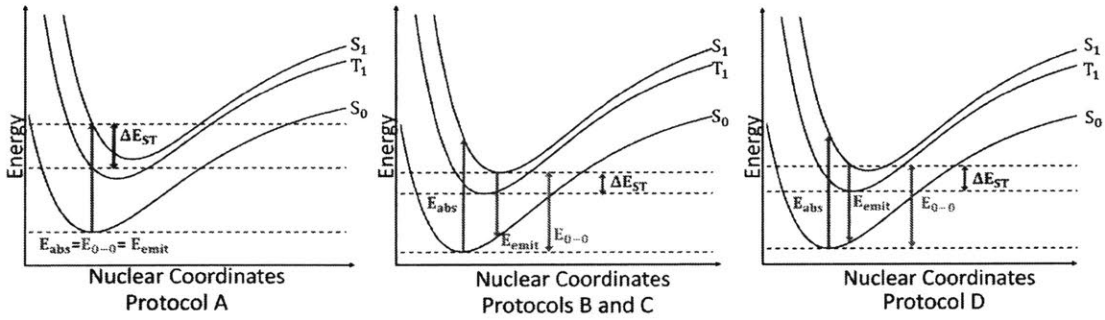


Figure 6-11: Photophysical parameters measured by the protocols. The arrows only indicate energy gaps and not nuclear coordinates of transitions.

The protocols themselves are as follows:

(1) Protocol A: S_0 geometry is optimized using ground-state DFT, and TDDFT is then employed to find the energies of the S_1 and T_1 states at this geometry. The resulting vertical absorption energy E_{abs} is then assumed to be a reasonable estimate for both E_{emit} and E_{0-0} (Figure 6-11, left panel). ΔE_{ST} is assumed to be the difference between S_1 and T_1 energies at the equilibrium S_0 geometry. This is the computationally cheapest of all the protocols, as it involves only one ground-state geometry optimization. However, it compromises the physics as real TADF molecules have nonzero experimental Stokes shifts and is not at all likely to be effective when the exact functional is employed.

(2) Protocol B: S_0 geometry is optimized using ground-state DFT, while both S_1 and T_1 geometries are optimized by TDDFT. TDDFT is then employed to find E_{abs} and E_{emit} as the vertical transition energy between the S_0 and S_1 surfaces, starting from the equilibrium

S_0 and S_1 geometries, respectively. E_{0-0} is obtained from the difference in the equilibrium S_1 and S_0 energies (found by TDDFT and ground-state DFT, respectively), and ΔE_{ST} is given by the difference between equilibrium S_1 and T_1 energies (found by TDDFT). Unlike Protocol A, this protocol does not compromise the physics, as the calculated parameters correspond exactly with experimentally measured ones.

(3) Protocol C: S_0 geometry is optimized by ground-state DFT, while the T_1 geometry is optimized with restricted open-shell DFT (RO-DFT). The equilibrium S_1 geometry is obtained via ROKS. The energy differences are then found in the same manner as Protocol B, except that ROKS and RO-DFT are used instead of TDDFT to calculate S_1 and T_1 energies, respectively. RO-DFT is preferred over unrestricted open-shell DFT for accessing T_1 energies in order to avoid systematic errors in ΔE_{ST}. Like Protocol B, the parameters calculated with this protocol also correspond exactly to experimentally measured ones.

(4) Protocol D: S_0 geometry is optimized by ground-state DFT, while the T_1 geometry is optimized with RO-DFT. It is assumed that the equilibrium S_1 geometry is fairly well approximated by the T_1 geometry (which is definitely the case for molecules with large CT character where ΔE_{ST} is small), and the energy differences are then found in the same manner as Protocol C. Overall, only two geometry optimizations (both of which were formally in the ground state) are employed, making this significantly cheaper than Protocol C.

The accuracy of the different protocols are compared by applying them to a set of relevant TADF chromophores, as shown in Figure 6-10. The experimental results are collected from work by Adachi et al [346, 364, 373, 374].

6.3.2 Performance on TADF test sets

The errors associated with Protocols A and B are given in Table 6.3, while the errors for Protocols C and D can be found in Table 6.4. Because of the wide spectrum of values for the experimental ΔE_{ST} associated with the test set, we report the errors in $log(\Delta E_{ST})$ instead of errors in ΔE_{ST}.

TDDFT results: Protocols A and B. Protocol A fares quite badly in estimating E_{abs} with all the functionals. This is unsurprising for semi-local PBE, B3LYP, and PBE0 calculations in light of the large CT nature of the S_1 state, which leads to systematic

Parameter measured	PBE RMSE (ME)	B3LYP RMSE (ME)	PBE0 RMSE (ME)	LC-ωPBE RMSE (ME)
E_{abs} (Protocols A and B)	1.38 (-1.31)	0.57 (-0.45)	0.43 (-0.28)	0.74 (0.72)
E_{emit} (Protocol A)	0.81 (-0.75)	0.25 (0.11)	0.35 (0.28)	1.29 (1.28)
E_{emit} (Protocol B)	1.48 (-1.45)	0.65 (-0.62)	0.49 (-0.44)	0.66 (0.63)
E_{0-0} (Protocol A)	1.00 (-0.95)	0.26 (-0.10)	0.23 (0.07)	1.04 (1.03)
E_{0-0} (Protocol B)	1.30 (-1.25)	0.52 (-0.46)	0.37 (-0.28)	0.82 (0.81)
$log(\Delta E_{ST})$ (Protocol A)	0.92 (-0.74)	0.68 (-0.37)	0.45 (-0.09)	0.76 (0.71)
$log(\Delta E_{ST})$ (Protocol B)	1.45 (-1.34)	1.04 (-0.85)	0.85 (-0.60)	0.74 (0.69)

Table 6.3: Errors associated with energy estimates from TDDFT-derived protocols. ME = mean error; RMSE = root mean squared error. Errors in E_{abs}, E_{emit}, and E_{0-0} have the unit eV. Both protocols calculate absorption energy in the same manner and thus have the same errors associated with that parameter.

underestimation of S_1 energy. Even the long-range corrected LC-ωPBE is not successful in estimating E_{abs}, although it systematically overestimates energy unlike the other three. This behavior indicates that the length scale of charge transfer for the test set molecules is smaller than ω^{-1} and is consistent with what was reported earlier by Adachi et al. [364]. However, the spurious TDDFT red-shift of energies in the B3LYP and PBE0 calculations are comparable to the Stokes shift for many of the molecules, resulting in a cancellation of errors that permits E_{abs}^{TDDFT} to be a fairly accurate estimate of E_{0-0}^{Expt}. With B3LYP, it is also possible to get reasonable estimates of E_{emit}^{Expt} by using E_{abs}^{TDDFT}, although PBE0 significantly overestimates this parameter (possibly on account of using a greater percentage of exact exchange than B3LYP). The energy shifts for PBE are too large for a similar cancellation of errors to occur there, and LC-ωPBE overestimates energies, making such a cancellation impossible.

This cancellation of errors however is not applicable to the ΔE_{ST} estimates, and Protocol A does not perform particularly well on that front. PBE and B3LYP both have large systematic errors as TDDFT artificially increases the extent of CT in the S_1 and T_1 states in a bid to lower their energies. This spuriously increased CT character leads to a smaller than expected energy gap, causing underestimation of ΔE_{ST}. PBE0 however has a smaller systematic error in ΔE_{ST}, possibly on account of the larger proportion of exact exchange being employed in the functional (a trend that can also been seen in the decreased errors on going from PBE to B3LYP). LC-ωPBE on the other hand, significantly overestimates ΔE_{ST},

which is consistent with the CT length scale being smaller than ω^{-1}. Consequently, single-functional Protocol A calculations cannot generally be used to get accurate ideas about ΔE_{ST} and E_{emit} simultaneously. The former is predicted best with PBE0 and the latter by B3LYP. However, Protocol A does not require excited-state geometry optimizations and is thus attractive as a preliminary screen for OLED materials, even if two calculations with different functionals are required.

Protocol B has no cancellation of errors to fall back upon, and thus consistently underestimates E_{emit} and E_{0-0} with PBE, B3LYP, and PBE0. The ΔE_{ST} estimates are also considerably underestimated (by nearly an order of magnitude) and are in fact much worse than Protocol A estimates. This is a consequence of TDDFT further enhancing the CT character of S_1/T_1 states by distorting the equilibrium geometry in an attempt to spuriously lower the energy. LC-ωPBE again overestimates parameters significantly, for the same reason as earlier [364]. Despite Protocol B being the most computationally expensive of the protocols tested (as it requires three geometry optimizations, two of which were in the excited state), it proves to be the least effective in predicting energies. This behavior is consistent with earlier studies and only serves to reinforce the notion that TDDFT with traditional functionals is unsuitable for predicting energies of CT states.

Parameter measured	PBE RMSE (ME)	B3LYP RMSE (ME)	PBE0 RMSE (ME)	LC-ωPBE RMSE (ME)
E_{abs} (Protocols C and D)	0.69 (-0.64)	0.18 (-0.06)	0.28 (0.11)	0.82 (0.75)
E_{emit} (Protocol C)	0.53 (-0.49)	0.19 (0.00)	0.20 (0.10)	0.79 (0.45)
E_{emit} (Protocol D)	0.52 (-0.48)	0.22 (0.02)	0.23 (0.11)	0.85 (0.79)
E_{0-0} (Protocol C)	0.56 (-0.54)	0.14 (0.02)	0.17 (0.11)	0.70 (0.66)
E_{0-0} (Protocol D)	0.55 (-0.52)	0.23 (0.10)	0.27 (0.19)	1.09 (1.03)
$log(\Delta E_{ST})$ (Protocol C)	0.73 (-0.55)	0.35 (-0.17)	0.27 (-0.07)	0.58 (-0.06)
$log(\Delta E_{ST})$ (Protocol D)	0.49 (-0.34)	0.35 (-0.03)	0.32 (0.04)	0.64 (0.25)

Table 6.4: Errors associated with energy estimates from ROKS-derived protocols. ME = mean error; RMSE = root mean squared error. Errors in E_{abs}, E_{emit}, and E_{0-0} have the unit eV. Both protocols calculate absorption energy in the same manner and thus have the same errors associated with that parameter.

ROKS results: Protocols C and D. Protocols C and D attempt to circumvent the red-shifting of energies by using ROKS and restricted open-shell DFT (RO-DFT) instead of

TDDFT to access the S_1 and T_1 surfaces, respectively. While it is possible to use unrestricted open-shell DFT (UO-DFT) to access T_1 energies instead, we choose to use RO-DFT as S_1 energies are always accessed by a restricted method (ROKS) and thus using UO-DFT triplet energies could lead to a systematic error in ΔE_{ST} from the extra stabilization recovered by the unrestricted calculation on the triplet. Nonetheless, we compared UO-DFT and RO-DFT calculation results for the case of the B3LYP functional and find that the numbers are not significantly different, further suggesting that performing RO-DFT is sufficient.

Though Protocol C is not a particularly cheap protocol (it requires three geometry optimizations, although only one of them is in the excited state), it gives quite accurate estimates of E_{abs}, E_{emit}, and E_{0-0} with B3LYP and PBE0. The near zero mean errors of B3LYP calculations are of particular interest, as they indicate there is no systematic bias unlike Protocol B with B3LYP. Interestingly, PBE0 mean errors are consistently greater than B3LYP errors by approximately 0.1 eV, possibly on account of PBE0 blue-shifting the energies more due to greater exact exchange. PBE still underestimates energies, but the deviation is still much less than Protocol B (by approximately 0.5 eV) or even Protocol A. LC-ωPBE still overestimates energies, and the deviations here are comparable to the TDDFT deviations, suggesting that these are more a consequence of the functional than the method.

The trends in ΔE_{ST} obtained from Protocol C are somewhat more interesting. B3LYP and PBE0 have the smallest RMS errors, which coupled with their relatively small mean errors indicate that these two are best suited for calculating ΔE_{ST} (PBE0 being somewhat better than B3LYP). Like in TDDFT calculations, PBE significantly underestimates ΔE_{ST}, although the errors are smaller. On the other hand, LC-ωPBE has an extremely small mean error along with a fairly large RMS error, indicating that a lot of noise is associated with calculations based on this functional but not much of a bias, which stands in direct contrast to the large Protocols A and B mean errors. This indicates that the ROKS/RO-DFT combination does not add a systematic bias to ΔE_{ST} for LC-ωPBE calculations, unlike TDDFT – although ΔE_{ST} calculations are still fairly inaccurate because this functional causes a large blue-shift of the S_1 and T_1 energies, which leads to a lot of noise. Overall, Protocol C is found to give very accurate energies with B3LYP and PBE0 and should be the method of choice if sufficient computational resources for S_1 geometry optimization are available. It is also possible to reduce the S_1 optimization cost by using T_1 optimized geometries as

the starting guess, as those are expected to be closer to equilibrium S_1 geometry than S_0 geometries or ground-state force field fits for TADF molecules.

Protocol D aims to attain accuracy comparable to Protocol C at a lesser computational cost by approximating the equilibrium S_1 geometry with the equilibrium T_1 geometry. This approach is reasonable for systems exhibiting TADF, as ΔE_{ST} is very small in these cases, indicating that the S_1 and T_1 surfaces are near parallel. Overall, Protocol D E_{emit} estimates are quite close to Protocol C estimates and thus correspond well to experimental values for B3LYP and PBE0 functionals. PBE calculations also give E_{0-0} similar to Protocol C, although the other three functionals overestimate this parameter relative to Protocol C (although the shift is only of the order of 0.1 eV on average for PBE0 and B3LYP).

This deficiency is somewhat compensated by the lack of apparent systematic bias in ΔE_{ST} calculated with PBE0 and B3LYP. PBE underestimates ΔE_{ST} as in all previous calculations, while LC-ωPBE overestimates ΔE_{ST} on average – unlike in Protocol C. Overall, all the functionals overestimate ΔE_{ST} with Protocol D, relative to Protocol C, as Protocol D slightly overestimates S_1 energy as it is evaluated close to but not at the minima. However, this slight blue-shift of ΔE_{ST} improves the mean accuracy of both PBE0 and B3LYP and allows Protocol D with these two functionals to be either better than or as effective as Protocol C in estimating E_{emit} and ΔE_{ST}, the two parameters of greatest interest for OLED screening. Protocol D with B3LYP/PBE0 is therefore the method we recommend be used first, with Protocol C only being used for the cases where D predicts large ΔE_{ST} ($>$ 0.5 eV), indicating smaller than expected CT character that causes our assumption about similarities in equilibrium S_1 and T_1 geometries to break down.

Overall, it can be seen that ROKS-derived protocols employing B3LYP or PBE0 yield numbers that are much closer to experimental values than the equivalent TDDFT-based protocols. We believe this is on account of ROKS computing the optimal orbitals for the excited-state self-consistently without any interference from the S_0 state (though S_0 orbitals serve as an initial guess). TDDFT on the other hand accesses excited states via linear response from the S_0 density, leading to the possibility of ground-state contamination when employed with approximate functionals despite TDDFT being guaranteed to be exact with the exact functional [63]. Such contamination would be especially problematic for CT states as these states are very far off from the S_0 state in terms of density, leading to a spurious lowering of energy that SCF methods like ROKS avoid by explicitly calculating the excited-

state density independent of the S_0 state. Similar reasons explain the lesser accuracy of T_1 energies obtained by TDDFT compared to formally exact SCF methods like RO-DFT and UO-DFT. The ROKS part of this analysis is functional dependent: TDDFT is guaranteed to give the exact answer with the exact functional, while no equivalent assurance exists for ROKS. However, ROKS appears to give a more accurate picture of CT states than TDDFT with commonly used semi-local hybrid density functionals like B3LYP and PBE0 at least, and both seem to be incorrect to a roughly equal extent with PBE or LC-ωPBE functionals.

6.4 Conclusions

In this chapter, we have shown successful applications of the conventional excited-state DFT method (TDDFT with B3LYP functional) for designing new TADF materials. Using two design strategies, we have designed the triptycene-based and TCZ-TRZ-based compounds, and shown they exhibit good TADF properties and OLED device performance. These results indicate that conventional TDDFT methods can serve as a useful computational design and screening tool as long as employed properly, although they also possess systematic errors. To address the problems existing in TDDFT, we have developed new computational protocols that are very well suited for studying molecules with large CT character, based on the ROKS/RO-DFT approach. Our numerical tests show that such ROKS-based protocols perform better than TDDFT on the TADF molecules. In conclusion, we have demonstrated how to choose suitable exicted-state DFT methods for studying photophysical properties of particular interest in this work, using the OLED emitters with TADF character as an example.

Our results however neglected the impact of the surroundings on the photophysics of these TADF molecules, mainly because the effects of the surrounding molecules are difficult to account for. This study only performed calculations on molecules that were experimentally studied in non-polar solvents like cyclohexane ($\epsilon_r = 2.03$) or toluene ($\epsilon_r = 2.39$), in part because the small dielectric constants should have a proportionately small effect on the photophysics. In the future, it would be interesting to examine how these protocols could be extended to deal accurately with the effects of the surroundings – for example, to tell the difference between the Stokes shift in solution versus in a film. We are also currently unable to predict quantum yields from first principles, which is another important param-

eter to be considered for practical applications. Our future work therefore shall focus on properly accounting for solvent effects on photophysical properties and *ab initio* quantum yield prediction in order to enable more efficient design of organic semiconductors involving CT states.

6.5 Acknowledgements

The experimental work of tryptycene-based TADF design is performed by Dr. Katsuaki Kawasumi and Dr. Tony Wu. The experimental work of TCZ-TRZ-based TADF design is performed by Dr. Wenliang Huang and Markus Einzinger. The work on application of ROKS to TADF emitters is carried out jointly with Diptarka Hait and Dr. David McMahon.

Chapter 7

Conclusions

In this thesis, we present our efforts on two aspects of DFT simulations in complex chemical systems: (1) building a systematically improvable density functional hierarchy, and (2) applying DFT methods to model electronic properties in a proper manner.

In Chapters 2-4, We have developed a new fragment-based density functional hierarchy, many-pair expansion (MPE), in order to systematically correct the deficiencies of any approximate density functional. We have shown that MPE at low orders can already accurately describe strong corrections presented in 1D/2D Hubbard and 1D Peierls-Hubbard lattice models and dispersion interactions presented in the PPP model, which are two big challenges for commonly-used density functionals. We have also extended MPE to molecular systems and demonstrated that MPE at a low second order provides accurate predictions for molecular and reaction energies of a series of small molecules and also correctly describes the bond breaking processes in hydrogen rings. More importantly, for most of the systems and problems tested, MPE is able to provide systematic improvement of approximate density functional calculations as successive higher order corrections are applied, as long as a suitable density partitioning method is prescribed. During the implementation of MPE in molecular systems, we have developed a new density partitioning method, self-attractive Hartree (SAH), which not only provides the foundation for molecular MPE calculations, but also shows promise as a tool to analyze chemical bonding from the electron density. Furthermore, we have shown that the SAH decomposition can be utilized to study monomer properties in molecular complexes and locally measure intra- and inter-molecular hydrogen bonding strength.

Systematic improvement of approximate density functionals is probably the biggest challenge in the field of DFT. The work in this thesis is encouraging as MPE shows promise to provide such a framework towards this goal. Nevertheless, we also note MPE needs further development to be finally employed in routine DFT simulations of materials. The future work of MPE should focus on four aspects.

First, more realistic density functionals should be used to test the performance of MPE. In this thesis, only the EXX functional is employed to test the performance of MPE in molecular systems. As correlation is completely missing in the EXX functional, the MPE convergence may be slow in systems where EXX is performing poorly. Combining MPE with more complicated functionals, it is possible that we can achieve better MPE convergence than EXX-MPE due to the introduction of correlation functionals. On the other hand, for modern functionals, especially empirically-tuned functionals such as B3LYP, it remains unclear if adding corrections in small fragments will ruin the original balance between the exchange and correlation functionals. Therefore, in addition to EXX, we should test whether MPE can also systematically improve the DFT calculations with commonly-used LDA, GGA, meta-GGA and hybrid functionals. Especially, we want to investigate how MPE can be applied to remove the delocalization error and strong correlation error in these functionals. Considering the elegance in MPE formalisms and the excellent performance of LDA-MPE in the 1D Hubbard model, we hope MPE can achieve better performance when combined with more realistic density functionals.

Second, faster implementation is needed for applying MPE to more realistic chemical systems. Although the MPEm correction is an $O(N^m)$ method (N is the number of electron pairs), its speed is usually hindered by the large prefactor due to expensive interacting potential inversion calculations. This limits MPE to be only applicable to small molecules and/or at low orders. To make MPE faster, two different ways may be explored. The first strategy is to combine DFT embedding methods [83, 84] with MPE to avoid doing interacting potential inversions. Instead of partitioning the density into pair densities in MPE, we may partition the external potential into "pair potentials" which give pair densities in the framework of DFT embedding. Starting from the partitioned potentials, the calculation of FCI energies does not require doing potential inversions any more. Our SAH decomposition method provides the possibility for such potential partitioning and we are actively exploring this new idea. The second strategy is to exploit the locality in SAH fragment densities and

using local correlation methods [207, 378, 379] as the alternative to FCI. In this thesis, all fragment calculations in molecular MPE use the basis set of the whole system, which significantly limits the application of MPE. However, it is entirely possible to adopt localization correlation methods to accelerate the time-consuming FCI calculations considering our SAH fragments are highly localized. Note that we have already exploited locality in the lattice model studies of MPE, where MPE at all orders can be seen as near-linear-scaling methods with different prefactors.

Third, a density optimization scheme is needed for more accurate MPE calculations. MPE is formally a density functional hierarchy that converges to the exact energy if the input density is exact. However, in reality, MPE results may converge to the inaccurate answer due to the use of bad approximate DFT density. One can already see such effects from the case of hydrogen ring bond breaking. Therefore, a self-consistent MPE scheme may be needed, where the MPE energies are used to optimize the input density and thus lead to better MPE results.

Fourth, MPE may be used as a tool to analyze the density functional errors and develop better density functionals. As shown in this work, MPE is able to reveal how approximate functionals perform for particular fragment densities by comparing with the exact correlation and interaction energies from FCI. Such analysis provides useful insights for functional development. New density functionals can thus be designed by optimally tuning parameters (e.g., range-separation parameter, percentage of EXX) for specific molecule so that certain fragment correction/interaction energies (e.g., pair-pair interactions) can be accurately described, leading to improved description of the whole molecule.

In Chapters 5-6, we have applied various DFT methods to study organic light-emitting diodes (OLEDs). We first show a condensed phase modeling of the emission layer of a phosphorescent OLED, consisting of guest and host materials. Using combined MD and QM/MM simulations, we have uncovered the charge recombination pathways and host exciton quenching mechanisms in the studied host-guest OLED. Based on the revealed mechanisms, we have proposed several strategies for designing better host and guest combinations. Then we utilize several excited-state DFT methods to computationally screen thermally activated delayed fluorescence (TADF) materials, which results in successful design of new efficient TADF emitters. In addition, we have developed a more accurate excited-state DFT protocols to predict TADF energetics based on a ROKS approach.

Using the example of OLED simulations, we have shown that DFT methods provide efficient modeling of electronic properties in complex chemical systems when employed properly. Incorporating environment effects and utilizing accurate excited-state methods are especially important for obtaining electronic properties in systems like organic semiconductors. Looking forward, several more developments should be considered. First, other electronic properties should be considered in OLED simulations. In this thesis, we have only modeled energy-related properties, while several other important electronic properties may also be needed. For example, for more comprehensive understanding of charge and energy transfer mechanisms, the electronic coupling between involving states should be considered. To more accurately predict TADF properties, an *ab initio* method to predict quantum yields is highly desired. Second, better methods to describe the environment effects are needed. In this thesis, we adopt a polarizable QM/MM method for the condensed phase simulation, where the classical MM force field is used to describe the environment. However, the classical force fields are usually not optimized for the interested organic semiconductor systems, which may introduce unexpected errors. Meanwhile, some parameters for particular chemical structures are missing in the force field, which requires extra efforts in fitting force field parameters to DFT or experimental results. Therefore, more accurate while still efficient first-principle methods may be needed to describe the environment effects. Subsystem DFT [225] and the recently-developed embedded mean-field theory [380] are possible methods to be explored.

Appendix A

Construction of auxiliary basis sets in SAH and MPE calculations

In Chapter 3, in order to implement self-attractive Hartree decomposition using standard linear algebra techniques, we need to have a way to represent the potential $\mu(\mathbf{r})$ in Eq. 3.8. We introduce an auxiliary basis set $\{\chi_P\}$ [226, 227] composed of Gaussian-type atomic orbitals. Then the external potential $\mu(\mathbf{r})$ can be expanded using the nuclear potential and Coulomb potentials of functions in the auxiliary basis set [81] as shown in Eqs. 3.9-3.14.

To solve Eq. 3.14, potential inversion techniques need to be exploited to guarantee density matching. As pointed out by many authors, potential inversion and optimized effective potential techniques are numerically unstable in finite basis sets [233, 234]. Görling and co-workers proposed a solution by carefully balancing the orbital ($\{\phi_i\}$) and auxiliary basis sets $\{\chi_P\}$ [235]. We follow their idea to decontract the orbital basis and remove some of the most compact and diffuse functions to generate auxiliary basis sets. In Chapter 4, we adopt the same orbital and auxiliary basis sets to perform potential inversions in MPE calculations.

Here we describe how to construct the orbital and auxiliary basis sets for SAH and MPE calculations. In this thesis, the orbital basis is constructed using uncontracted cc-pVDZ, cc-pVTZ [231] or aug-cc-pVTZ basis set [232]. The corresponding auxiliary basis sets are listed in Table SA.1 for elements including H, Be, B, C, N, O, F, S and Se.

Element	Orbital Basis	Auxiliary basis			
			Removed		Added
		L	Tight	Diffuse	Exponent
H	cc-pVDZ	0	1	1	
	cc-pVTZ	0	1	1	
		1		1	
	aug-cc-pVTZ	0	1		
Be	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	
		1		1	
		3		1	
B	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	
		1		1	
		2		1	0.2109
		3		1	
C	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	
		1		1	
		2		1	0.3499
		3		1	
	aug-cc-pVTZ	0	3	1	
		1		1	
		2		1	
		3	1	1	
N	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	

		1		1	
		3		1	
	aug-cc-pVTZ	0	3	1	
		1		1	
		2		1	
		3	1	1	
O	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	
		1		1	
		2		1	0.7163
		3		1	
	aug-cc-pVTZ	0	3	1	
		1		1	
		2		1	
		3	1	1	
F	cc-pVDZ	0	3	1	
		1		1	
	cc-pVTZ	0	3	1	
		1		1	
		2		1	0.9684
		3		1	
	aug-cc-pVTZ	0	3	1	
		1		1	
		2		1	
		3	1	1	
S	cc-pVTZ	0	4	1	
		1	1	1	
		2		1	
		3		1	
Se	cc-pVTZ	0	4	1	
		1	1	1	

		2		1	
		3		1	

Table A.1: The construction of auxiliary basis sets for elements used in this thesis. The auxiliary basis set is designed for specific orbital basis by removing several basis functions from uncontracted orbital basis. The second column shows the reference uncontracted orbital basis for generating auxiliary basis. The third column shows the angular quantum number of removed/added Gaussian functions. The fourth and fifth column shows the number of most compact and diffuse Gaussian functions removed from the uncontracted orbital basis. The last column shows the exponent of the added Gaussian functions in the auxiliary basis.

Bibliography

[1] I. N. Levine, D. H. Busch, and H. Shull, *Quantum chemistry*, vol. 5. Prentice Hall Upper Saddle River, NJ, 2000.

[2] E. Schrödinger, "An undulatory theory of the mechanics of atoms and molecules," *Phys. Rev.*, vol. 28, no. 6, p. 1049, 1926.

[3] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Intro to Advanced Electronic Structure Theory*. Dover publications, 1996.

[4] J. R. Thomas, B. J. DeLeeuw, G. Vacek, T. D. Crawford, Y. Yamaguchi, and H. F. Schaefer III, "The balance between theoretical method and basis set quality: A systematic study of equilibrium geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities," *J. Chem. Phys.*, vol. 99, no. 1, pp. 403–416, 1993.

[5] J. M. Martin and G. de Oliveira, "Towards standard methods for benchmark quality ab initio thermochemistry – w1 and w2 theory," *J. Chem. Phys.*, vol. 111, no. 5, pp. 1843–1856, 1999.

[6] Y. Zhao, N. González-García, and D. G. Truhlar, "Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods," *J. Phys. Chem. A*, vol. 109, no. 9, pp. 2012–2018, 2005.

[7] J. F. Stanton and R. J. Bartlett, "The equation of motion coupled-cluster method. a systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties," *J. Chem. Phys.*, vol. 98, no. 9, pp. 7029–7039, 1993.

[8] C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, "Natural triple excitations in local coupled cluster calculations with pair natural orbitals," *J. Chem. Phys.*, vol. 139, no. 13, p. 134101, 2013.

[9] R. Z. Khaliullin, J. VandeVondele, and J. Hutter, "Efficient linear-scaling density functional theory for molecular systems," *J. Chem. Theory Comput.*, vol. 9, no. 10, pp. 4421–4427, 2013.

[10] R. A. Friesner, "Ab initio quantum chemistry: Methodology and applications," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 102, no. 19, pp. 6648–6653, 2005.

[11] J. K. Nørskov, F. Abild-Pedersen, F. Studt, and T. Bligaard, "Density functional theory in surface chemistry and catalysis," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 108, no. 3, pp. 937–943, 2011.

[12] M. G. Mavros, T. Tsuchimochi, T. Kowalczyk, A. McIsaac, L.-P. Wang, and T. V. Voorhis, "What can density functional theory tell us about artificial catalytic water splitting?," *Inorg. Chem.*, vol. 53, no. 13, pp. 6386–6397, 2014.

[13] V. Guallar and R. A. Friesner, "Cytochrome p450cam enzymatic catalysis cycle: a quantum mechanics/molecular mechanics study," *J. Am. Chem. Soc.*, vol. 126, no. 27, pp. 8501–8508, 2004.

[14] S. R. Yost, J. Lee, M. W. Wilson, T. Wu, D. P. McMahon, R. R. Parkhurst, N. J. Thompson, D. N. Congreve, A. Rao, K. Johnson, M. Y. Sfeir, M. G. Bawendi, T. M. Swager, R. H. Friend, M. A. Baldo, and T. Van Voorhis, "A transferable model for singlet-fission kinetics," *Nat. Chem.*, vol. 6, no. 6, p. 492, 2014.

[15] F. De Angelis, S. Fantacci, and A. Selloni, "Alignment of the dye's molecular levels with the tio2 band edges in dye-sensitized solar cells: a dft–tddft study," *Nanotechnology*, vol. 19, no. 42, p. 424002, 2008.

[16] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, and A. Aspuru-Guzik, "Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach," *Nat. Mater.*, vol. 15, no. 10, p. 1120, 2016.

[17] H. Sun, C. Zhong, and J.-L. Bredas, "Reliable prediction with tuned range-separated functionals of the singlet–triplet gap in organic emitters for thermally activated delayed fluorescence," *J. Chem. Theory Comput.*, vol. 11, no. 8, pp. 3851–3858, 2015.

[18] C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. Van Voorhis, and B. L. Pentelute, "π-clamp-mediated cysteine conjugation," *Nat. Chem.*, vol. 8, no. 2, p. 120, 2016.

[19] P. Dai, C. Zhang, M. Welborn, J. J. Shepherd, T. Zhu, T. Van Voorhis, and B. L. Pentelute, "Salt effect accelerates site-selective cysteine bioconjugation," *ACS Cent. Sci.*, vol. 2, no. 9, pp. 637–646, 2016.

[20] Q. Cui and M. Karplus, "Triosephosphate isomerase: a theoretical comparison of alternative pathways," *J. Am. Chem. Soc.*, vol. 123, no. 10, pp. 2284–2290, 2001.

[21] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, no. 4A, p. A1133, 1965.

[22] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules.* Oxford University Press, New York, 1989.

[23] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Challenges for density functional theory," *Chem. Rev.*, vol. 112, no. 1, pp. 289–320, 2011.

[24] B. J. Lynch, P. L. Fast, M. Harris, and D. G. Truhlar, "Adiabatic connection for kinetics," *J. Phys. Chem. A*, vol. 104, no. 21, pp. 4811–4815, 2000.

[25] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," *Phys. Rev. Lett.*, vol. 100, no. 13, p. 136406, 2008.

[26] Y. Zhao and D. 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. Acc.*, vol. 120, no. 1-3, pp. 215–241, 2008.

[27] C. W. Tang and S. A. VanSlyke, "Organic electroluminescent diodes," *Appl. Phys. Lett.*, vol. 51, no. 12, pp. 913–915, 1987.

[28] H. Uoyama, K. Goushi, K. Shizu, H. Nomura, and C. Adachi, "Highly efficient organic light-emitting diodes from delayed fluorescence," *Nature*, vol. 492, no. 7428, pp. 234–238, 2012.

[29] M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln," *Ann. Phys. (Berl.)*, vol. 389, no. 20, pp. 457–484, 1927.

[30] P. A. Dirac, "Quantum mechanics of many-electron systems," *Proc. R. Soc. Lond. A*, vol. 123, no. 792, pp. 714–733, 1929.

[31] P. W. Atkins and R. S. Friedman, *Molecular quantum mechanics*. Oxford university press, 2011.

[32] J. C. Slater, "The theory of complex spectra," *Phys. Rev.*, vol. 34, no. 10, p. 1293, 1929.

[33] C. D. Sherrill and H. F. Schaefer III, "The configuration interaction method: Advances in highly correlated approaches," in *Advances in quantum chemistry*, vol. 34, pp. 143–269, Elsevier, 1999.

[34] C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," *Phys. Rev.*, vol. 46, no. 7, p. 618, 1934.

[35] J. Čížek, "On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods," *J. Chem. Phys.*, vol. 45, no. 11, pp. 4256–4266, 1966.

[36] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories," *Chem. Phys. Lett.*, vol. 157, no. 6, pp. 479–483, 1989.

[37] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, no. 3B, p. B864, 1964.

[38] L. H. Thomas, "The calculation of atomic fields," in *Math. Proc. Camb. Philos. Soc.*, vol. 23, pp. 542–548, Cambridge University Press, 1927.

[39] E. Fermi, "Un metodo statistico per la determinazione di alcune priorieta dell'atome," *Rend. Accad. Naz. Lincei*, vol. 6, no. 602-607, p. 32, 1927.

[40] J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy," in *AIP Conf. Proc.*, vol. 577, pp. 1–20, AIP, 2001.

[41] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, "Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits," *J. Chem. Phys.*, vol. 123, no. 6, p. 062201, 2005.

[42] D. R. Hartree, *The calculation of atomic structures.* Wiley, 1957.

[43] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, no. 18, p. 3865, 1996.

[44] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A*, vol. 38, no. 6, p. 3098, 1988.

[45] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, no. 2, p. 785, 1988.

[46] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation," *Phys. Rev. B*, vol. 46, no. 11, p. 6671, 1992.

[47] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids," *Phys. Rev. Lett.*, vol. 91, no. 14, p. 146401, 2003.

[48] J. P. Perdew, M. Ernzerhof, and K. Burke, "Rationale for mixing exact exchange with density functional approximations," *J. Chem. Phys.*, vol. 105, no. 22, pp. 9982–9985, 1996.

[49] A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange," *J. Chem. Phys.*, vol. 98, no. 7, pp. 5648–5652, 1993.

[50] P. Stephens, F. Devlin, C. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," *J. Phys. Chem.*, vol. 98, no. 45, pp. 11623–11627, 1994.

[51] A. Görling and M. Levy, "Exact kohn-sham scheme based on perturbation theory," *Phys. Rev. A*, vol. 50, no. 1, p. 196, 1994.

[52] T. Schwabe and S. Grimme, "Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability," *Phys. Chem. Chem. Phys.*, vol. 9, no. 26, pp. 3397–3406, 2007.

[53] Y. Zhang, X. Xu, and W. A. Goddard, "Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 13, pp. 4963–4968, 2009.

[54] B. G. Janesko, T. M. Henderson, and G. E. Scuseria, "Long-range-corrected hybrid density functionals including random phase approximation correlation: Application to noncovalent interactions," *J. Chem. Phys.*, vol. 131, no. 3, p. 034110, 2009.

[55] A. Grüneis, M. Marsman, J. Harl, L. Schimka, and G. Kresse, "Making the random phase approximation to electronic correlation accurate," *J. Chem. Phys.*, vol. 131, no. 15, p. 154115, 2009.

[56] F. Furche, "Molecular tests of the random phase approximation to the exchange-correlation energy functional," *Phys. Rev. B*, vol. 64, no. 19, p. 195120, 2001.

[57] F. Furche and T. Van Voorhis, "Fluctuation-dissipation theorem density-functional theory," *J. Chem. Phys.*, vol. 122, no. 16, p. 164106, 2005.

[58] M. Filatov and S. Shaik, "A spin-restricted ensemble-referenced kohn–sham method and its application to diradicaloid situations," *Chem. Phys. Lett.*, vol. 304, no. 5-6, pp. 429–437, 1999.

[59] T. Kowalczyk, T. Tsuchimochi, P.-T. Chen, L. Top, and T. Van Voorhis, "Excitation energies and stokes shifts from a restricted open-shell kohn-sham approach," *J. Chem. Phys.*, vol. 138, no. 16, p. 164101, 2013.

[60] D. Braga and G. Horowitz, "High-performance organic field-effect transistors," *Adv. Mater.*, vol. 21, no. 14-15, pp. 1473–1486, 2009.

[61] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, "Second-order perturbation theory with a casscf reference function," *J. Phys. Chem.*, vol. 94, no. 14, pp. 5483–5488, 1990.

[62] M. Schreiber, M. R. Silva-Junior, S. P. Sauer, and W. Thiel, "Benchmarks for electronically excited states: Caspt2, cc2, ccsd, and cc3," *J. Chem. Phys.*, vol. 128, no. 13, p. 134110, 2008.

[63] E. Runge and E. K. Gross, "Density-functional theory for time-dependent systems," *Phys. Rev. Lett.*, vol. 52, no. 12, p. 997, 1984.

[64] M. A. Marques and E. K. Gross, "Time-dependent density functional theory," *Annu. Rev. Phys. Chem.*, vol. 55, pp. 427–455, 2004.

[65] Q. Wu and T. Van Voorhis, "Direct optimization method to study constrained systems within density-functional theory," *Phys. Rev. A*, vol. 72, no. 2, p. 024502, 2005.

[66] Q. Wu and T. Van Voorhis, "Direct calculation of electron transfer parameters through constrained density functional theory," *J. Phys. Chem. A*, vol. 110, no. 29, pp. 9212–9218, 2006.

[67] B. Kaduk, T. Kowalczyk, and T. Van Voorhis, "Constrained density functional theory," *Chem. Rev.*, vol. 112, no. 1, pp. 321–370, 2011.

[68] R. A. Marcus, "On the theory of oxidation-reduction reactions involving electron transfer. i," *J. Chem. Phys.*, vol. 24, no. 5, pp. 966–978, 1956.

[69] R. A. Marcus, "Electron transfer reactions in chemistry. theory and experiment," *Rev. Mod. Phys.*, vol. 65, no. 3, p. 599, 1993.

[70] C. A. Mead and D. G. Truhlar, "Conditions for the definition of a strictly diabatic electronic basis for molecular systems," *J. Chem. Phys.*, vol. 77, no. 12, pp. 6090–6098, 1982.

[71] T. Van Voorhis, T. Kowalczyk, B. Kaduk, L.-P. Wang, C.-L. Cheng, and Q. Wu, "The diabatic picture of electron transfer, reaction barriers, and molecular dynamics," *Annu. Rev. Phys. Chem.*, vol. 61, pp. 149–170, 2010.

[72] A. D. Becke, "A multicenter numerical integration scheme for polyatomic molecules," *J. Chem. Phys.*, vol. 88, no. 4, pp. 2547–2553, 1988.

[73] F. L. Hirshfeld, "Bonded-atom fragments for describing molecular charge densities," *Theor. Chem. Acc.*, vol. 44, no. 2, pp. 129–138, 1977.

[74] M. B. Goldey, N. P. Brawand, M. Voros, and G. Galli, "Charge transport in nanostructured materials: Implementation and verification of constrained density functional theory," *J. Chem. Theory Comput.*, vol. 13, no. 6, pp. 2581–2590, 2017.

[75] Q. Wu and T. Van Voorhis, "Constrained density functional theory and its application in long-range electron transfer," *J. Chem. Theory Comput.*, vol. 2, no. 3, pp. 765–774, 2006.

[76] Q. Zhao, R. C. Morrison, and R. G. Parr, "From electron densities to kohn-sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies," *Phys. Rev. A*, vol. 50, no. 3, p. 2138, 1994.

[77] Q. Wu and W. Yang, "A direct optimization method for calculating density functionals and exchange–correlation potentials from electron densities," *J. Chem. Phys.*, vol. 118, no. 6, pp. 2498–2509, 2003.

[78] R. Van Leeuwen and E. Baerends, "Exchange-correlation potential with correct asymptotic behavior," *Phys. Rev. A*, vol. 49, no. 4, p. 2421, 1994.

[79] W. Yang and Q. Wu, "Direct method for optimized effective potentials in density-functional theory," *Phys. Rev. Lett.*, vol. 89, no. 14, p. 143002, 2002.

[80] S. Ivanov, S. Hirata, and R. J. Bartlett, "Exact exchange treatment for molecules in finite-basis-set kohn-sham theory," *Phys. Rev. Lett.*, vol. 83, no. 26, p. 5455, 1999.

[81] A. Görling, "New ks method for molecules based on an exchange charge density generating the exact local ks exchange potential," *Phys. Rev. Lett.*, vol. 83, no. 26, p. 5459, 1999.

[82] I. G. Ryabinkin, A. A. Kananenka, and V. N. Staroverov, "Accurate and efficient approximation to the optimized effective potential for exchange," *Phys. Rev. Lett.*, vol. 111, no. 1, p. 013001, 2013.

[83] J. D. Goodpaster, T. A. Barnes, F. R. Manby, and T. F. Miller III, "Density functional theory embedding for correlated wavefunctions: Improved methods for open-shell systems and transition metal complexes," *J. Chem. Phys.*, vol. 137, no. 22, p. 224113, 2012.

[84] C. Huang, M. Pavone, and E. A. Carter, "Quantum mechanical embedding theory based on a unique embedding potential," *J. Chem. Phys.*, vol. 134, no. 15, p. 154110, 2011.

[85] P. Elliott, K. Burke, M. H. Cohen, and A. Wasserman, "Partition density-functional theory," *Phys. Rev. A*, vol. 82, no. 2, p. 024501, 2010.

[86] Q. Wu and W. Yang, "Algebraic equation and iterative optimization for the optimized effective potential in density functional theory," *J. Theor. Comput. Chem.*, vol. 2, no. 04, pp. 627–638, 2003.

[87] M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 76, no. 12, pp. 6062–6065, 1979.

[88] C. J. Cramer and D. G. Truhlar, "Implicit solvation models: equilibria, structure, spectra, and dynamics," *Chem. Rev.*, vol. 99, no. 8, pp. 2161–2200, 1999.

[89] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chem. Rev.*, vol. 105, no. 8, pp. 2999–3094, 2005.

[90] O. Tapia and O. Goscinski, "Self-consistent reaction field theory of solvent effects," *Mol. Phys.*, vol. 29, no. 6, pp. 1653–1661, 1975.

[91] S. Miertuš, E. Scrocco, and J. Tomasi, "Electrostatic interaction of a solute with a continuum. a direct utilizaion of ab initio molecular potentials for the prevision of solvent effects," *Chem. Phys.*, vol. 55, no. 1, pp. 117–129, 1981.

[92] B. Mennucci, "Polarizable continuum model," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 2, no. 3, pp. 386–404, 2012.

[93] A. Klamt and G. Schüürmann, "Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient," *J. Chem. Soc., Perkin Trans. 2*, no. 5, pp. 799–805, 1993.

[94] A. Klamt, "The cosmo and cosmo-rs solvation models," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 8, no. 1, 2018.

[95] C. P. Kelly, C. J. Cramer, and D. G. Truhlar, "Sm6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute- water clusters," *J. Chem. Theory Comput.*, vol. 1, no. 6, pp. 1133–1152, 2005.

[96] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions," *J. Phys. Chem. B*, vol. 113, no. 18, pp. 6378–6396, 2009.

[97] M. Aguilar, "Separation of the electric polarization into fast and slow components: A comparison of two partition schemes," *J. Phys. Chem. A*, vol. 105, no. 45, pp. 10393–10396, 2001.

[98] A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," *J. Mol. Bio.*, vol. 103, no. 2, pp. 227–249, 1976.

[99] M. J. Field, P. A. Bash, and M. Karplus, "A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations," *J. Comput. Chem.*, vol. 11, no. 6, pp. 700–733, 1990.

[100] H. Lin and D. G. Truhlar, "Qm/mm: what have we learned, where are we, and where do we go from here?," *Theor. Chem. Acc.*, vol. 117, no. 2, p. 185, 2007.

[101] G. Lamoureux and B. Roux, "Modeling induced polarization with classical drude oscillators: Theory and molecular dynamics simulation algorithm," *J. Chem. Phys.*, vol. 119, no. 6, pp. 3025–3039, 2003.

[102] A. Stone, "Distributed multipole analysis, or how to describe a molecular charge distribution," *Chem. Phys. Lett.*, vol. 83, no. 2, pp. 233–239, 1981.

[103] W. Kohn, "Nobel lecture: Electronic structure of matter-wave functions and density functionals," *Rev. Mod. Phys.*, vol. 71, no. 5, p. 1253, 1999.

[104] K. Burke, "Perspective on density functional theory," *J. Chem. Phys.*, vol. 136, no. 15, p. 150901, 2012.

[105] A. D. Becke, "Perspective: Fifty years of density-functional theory in chemical physics," *J. Chem. Phys.*, vol. 140, no. 18, p. 18A301, 2014.

[106] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Phys. Rev. B*, vol. 23, no. 10, p. 5048, 1981.

[107] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, "Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals," *J. Chem. Phys.*, vol. 125, no. 19, p. 194112, 2006.

[108] P. Mori-Sánchez, A. J. Cohen, and W. Yang, "Localization and delocalization errors in density functional theory and implications for band-gap prediction," *Phys. Rev. Lett.*, vol. 100, no. 14, p. 146401, 2008.

[109] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Fractional spins and static correlation error in density functional theory," *J.Chem. Phys.*, vol. 129, no. 12, p. 121104, 2008.

[110] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Insights into current limitations of density functional theory," *Science*, vol. 321, no. 5890, pp. 792–794, 2008.

[111] C. D. Sherrill, "Frontiers in electronic structure theory," *J. Chem. Phys.*, vol. 132, no. 11, p. 110902, 2010.

[112] S. Kristyán and P. Pulay, "Can (semi) local density functional theory account for the london dispersion forces?," *Chem. Phys. Lett.*, vol. 229, no. 3, pp. 175–180, 1994.

[113] M. J. Allen and D. J. Tozer, "Helium dimer dispersion forces and correlation potentials in density functional theory," *J. Chem. Phys.*, vol. 117, no. 24, pp. 11113–11120, 2002.

[114] O. A. Vydrov and G. E. Scuseria, "Assessment of a long-range corrected hybrid functional," *J. Chem. Phys.*, vol. 125, no. 23, p. 234109, 2006.

[115] O. A. Vydrov and G. E. Scuseria, "Effect of the perdew–zunger self-interaction correction on the thermochemical performance of approximate density functionals," *J. Chem. Phys.*, vol. 121, no. 17, pp. 8187–8193, 2004.

[116] C. Li, X. Zheng, N. Q. Su, and W. Yang, "Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations," *Natl. Sci. Rev.*, 2017.

[117] V. I. Anisimov, J. Zaanen, and O. K. Andersen, "Band theory and mott insulators: Hubbard u instead of stoner i," *Phys. Rev. B*, vol. 44, no. 3, p. 943, 1991.

[118] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. Marianetti, "Electronic structure calculations with dynamical mean-field theory," *Rev. Mod. Phys.*, vol. 78, no. 3, p. 865, 2006.

[119] G. Li Manni, R. K. Carlson, S. Luo, D. Ma, J. Olsen, D. G. Truhlar, and L. Gagliardi, "Multiconfiguration pair-density functional theory," *J. Chem. Theory Comput.*, vol. 10, no. 9, pp. 3669–3680, 2014.

[120] F. Malet and P. Gori-Giorgi, "Strong correlation in kohn-sham density functional theory," *Phys. Rev. Lett.*, vol. 109, no. 24, p. 246402, 2012.

[121] E. Stoudenmire, L. O. Wagner, S. R. White, and K. Burke, "One-dimensional continuum electronic structure with the density-matrix renormalization group and its implications for density-functional theory," *Phys. Rev. Lett.*, vol. 109, no. 5, p. 056402, 2012.

[122] Q. Wu and W. Yang, "Empirical correction to density functional theory for van der waals interactions," *J. Chem. Phys.*, vol. 116, no. 2, pp. 515–524, 2002.

[123] S. Grimme, "Semiempirical gga-type density functional constructed with a long-range dispersion correction," *J. Comput. Chem.*, vol. 27, no. 15, pp. 1787–1799, 2006.

[124] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu," *J. Chem. Phys.*, vol. 132, no. 15, p. 154104, 2010.

[125] A. D. Becke and E. R. Johnson, "A density-functional model of the dispersion interaction," *J. Chem. Phys.*, vol. 123, no. 15, p. 154101, 2005.

[126] A. Tkatchenko, R. A. DiStasio Jr, R. Car, and M. Scheffler, "Accurate and efficient method for many-body van der waals interactions," *Phys. Rev. Lett.*, vol. 108, no. 23, p. 236402, 2012.

[127] S. N. Steinmann and C. Corminboeuf, "Comprehensive benchmarking of a density-dependent dispersion correction," *J. Chem. Theory Comput.*, vol. 7, no. 11, pp. 3567–3577, 2011.

[128] J. Tao, J. P. Perdew, and A. Ruzsinszky, "Accurate van der waals coefficients from density functional theory," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 109, no. 1, pp. 18–21, 2012.

[129] V. Osinga, S. Van Gisbergen, J. Snijders, and E. Baerends, "Density functional results for isotropic and anisotropic multipole polarizabilities and c6, c7, and c8 van der waals dispersion coefficients for molecules," *J. Chem. Phys.*, vol. 106, no. 12, pp. 5091–5101, 1997.

[130] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der waals density functional for general geometries," *Phys. Rev. Lett.*, vol. 92, no. 24, p. 246401, 2004.

[131] G. Román-Pérez and J. M. Soler, "Efficient implementation of a van der waals density functional: application to double-wall carbon nanotubes," *Phys. Rev. Lett.*, vol. 103, no. 9, p. 096102, 2009.

[132] O. A. Vydrov and T. Van Voorhis, "Nonlocal van der waals density functional made simple," *Phys. Rev. Lett.*, vol. 103, no. 6, p. 063004, 2009.

[133] O. A. Vydrov and T. Van Voorhis, "Nonlocal van der waals density functional: The simpler the better," *J. Chem. Phys.*, vol. 133, no. 24, p. 244103, 2010.

[134] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, "Higher-accuracy van der waals density functional," *Phys. Rev. B*, vol. 82, no. 8, p. 081101, 2010.

[135] L. Goerigk and S. Grimme, "Efficient and accurate double-hybrid-meta-gga density functionals – evaluation with the extended gmtkn30 database for general main group thermochemistry, kinetics, and noncovalent interactions," *J. Chem. Theory Comput.*, vol. 7, no. 2, pp. 291–309, 2010.

[136] D. C. Langreth and J. P. Perdew, "The exchange-correlation energy of a metallic surface," *Solid State Commun.*, vol. 17, no. 11, pp. 1425–1429, 1975.

[137] D. C. Langreth and J. P. Perdew, "Exchange-correlation energy of a metallic surface: Wave-vector analysis," *Phys. Rev. B*, vol. 15, no. 6, p. 2884, 1977.

[138] O. A. Vydrov and G. E. Scuseria, "Ionization potentials and electron affinities in the perdew–zunger self-interaction corrected density-functional theory," *J. Chem. Phys.*, vol. 122, no. 18, p. 184107, 2005.

[139] T. Zhu, P. de Silva, H. van Aggelen, and T. Van Voorhis, "Many-electron expansion: A density functional hierarchy for strongly correlated systems," *Phys. Rev. B*, vol. 93, no. 20, p. 201108, 2016.

[140] J. Hubbard, "Electron correlations in narrow energy bands," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 276, pp. 238–257, The Royal Society, 1963.

[141] R. E. Peierls, *Quantum theory of solids*. No. 23, Oxford University Press, 1955.

[142] R. Pariser and R. G. Parr, "A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. i.," *J. Chem. Phys.*, vol. 21, no. 3, pp. 466–471, 1953.

[143] R. Pariser and R. G. Parr, "A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. ii," *J. Chem. Phys.*, vol. 21, no. 5, pp. 767–776, 1953.

[144] J. Pople, "Electron interaction in unsaturated hydrocarbons," *Trans. Faraday Soc.*, vol. 49, pp. 1375–1385, 1953.

[145] P. de Silva, T. Zhu, and T. Van Voorhis, "Long-range interactions from the many-pair expansion: A different avenue to dispersion in dft," *J. Chem. Phys.*, vol. 146, no. 2, p. 024111, 2017.

[146] E. H. Lieb, "Density functionals for coulomb systems," in *Inequalities*, pp. 269–303, Springer, 2002.

[147] G. L. Bendazzoli and S. Evangelisti, "Computation and analysis of the full configuration interaction wave function of some simple systems," *Int. J. of Quant. Chem.*, vol. 48, no. S27, pp. 287–301, 1993.

[148] G. L. Bendazzoli and S. Evangelisti, "A vector and parallel full configuration interaction algorithm," *J. Chem. Phys.*, vol. 98, no. 4, pp. 3141–3150, 1993.

[149] S. Kümmel and L. Kronik, "Orbital-dependent density functionals: Theory and applications," *Rev. Mod. Phys.*, vol. 80, no. 1, p. 3, 2008.

[150] J. Cui, H. Liu, and K. D. Jordan, "Theoretical characterization of the (h2o) 21 cluster: Application of an n-body decomposition procedure," *J. Phys. Chem. B*, vol. 110, no. 38, pp. 18872–18878, 2006.

[151] R. M. Richard and J. M. Herbert, "A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory," *J. Chem. Phys.*, vol. 137, no. 6, p. 064113, 2012.

[152] H. Stoll, "On the correlation energy of graphite," *J. Chem. Phys.*, vol. 97, no. 11, pp. 8449–8454, 1992.

[153] L. Hedin, "New method for calculating the one-particle green's function with application to the electron-gas problem," *Phys. Rev.*, vol. 139, no. 3A, p. A796, 1965.

[154] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge university press, 2007.

[155] L. O. Wagner, T. E. Baker, E. Stoudenmire, K. Burke, and S. R. White, "Kohn-sham calculations with the exact functional," *Phys. Rev. B*, vol. 90, no. 4, p. 045109, 2014.

[156] W. Kohn, "v-representability and density functional theory," *Phys. Rev. Lett.*, vol. 51, no. 17, p. 1596, 1983.

[157] A. Savin and T. A. Wesolowski, "Orbital-free embedding effective potential in analytically solvable cases," in *Advances in the Theory of Atomic and Molecular Systems*, pp. 311–326, Springer, 2009.

[158] P. de Silva and T. A. Wesolowski, "Exact non-additive kinetic potentials in realistic chemical systems," *J. Chem. Phys.*, vol. 137, no. 9, p. 094110, 2012.

[159] P. de Silva and T. A. Wesolowski, "Pure-state noninteracting v-representability of electron densities from kohn-sham calculations with finite basis sets," *Phys. Rev. A*, vol. 85, no. 3, p. 032518, 2012.

[160] S. Klüpfel, P. Klüpfel, and H. Jónsson, "Importance of complex orbitals in calculating the self-interaction-corrected ground state of atoms," *Phys. Rev. A*, vol. 84, no. 5, p. 050501, 2011.

[161] D. Hofmann, S. Klüpfel, P. Klüpfel, and S. Kümmel, "Using complex degrees of freedom in the kohn-sham self-interaction correction," *Phys. Rev. A*, vol. 85, no. 6, p. 062514, 2012.

[162] S. Lehtola and H. JoÌĄnsson, "Variational, self-consistent implementation of the perdew–zunger self-interaction correction with complex optimal orbitals," *J. Chem. Theory Comput.*, vol. 10, no. 12, pp. 5324–5337, 2014.

[163] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, "General atomic and molecular electronic structure system," *J. Comput. Chem.*, vol. 14, no. 11, pp. 1347–1363, 1993.

[164] T. Schmidt, E. Kraisler, L. Kronik, and S. Kummel, "One-electron self-interaction and the asymptotics of the kohn-sham potential: an impaired relation," *Phys. Chem. Chem. Phys.*, vol. 16, pp. 14357–14367, 2014.

[165] P. de Silva and C. Corminboeuf, "Local hybrid functionals with orbital-free mixing function and balanced elimination o self-interaction error," *J. Chem. Phys.*, vol. 142, no. 7, p. 074112, 2015.

[166] P. de Silva and C. Corminboeuf, "Communication: A new class of non-empirical explicit density functionals on the third rung of jacob's ladder," *J. Chem. Phys.*, vol. 143, no. 11, p. 111105, 2015.

[167] P. de Silva and C. Corminboeuf, "Simultaneous visualization of covalent and noncovalent interactions using regions of density overlap," *J. Chem. Theory Comput.*, vol. 10, no. 9, pp. 3745–3756, 2014.

[168] J. Gräfenstein, E. Kraka, and D. Cremer, "The impact of the self-interaction error on the density functional theory description of dissociating radical cations: Ionic and covalent dissociation limits," *J. Chem. Phys.*, vol. 120, no. 2, pp. 524–539, 2004.

[169] P. Mori-Sánchez, A. J. Cohen, and W. Yang, "Many-electron self-interaction error in approximate density functionals," *J. Chem. Phys.*, vol. 125, no. 20, p. 201102, 2006.

[170] G. Knizia and G. K.-L. Chan, "Density matrix embedding: A simple alternative to dynamical mean-field theory," *Phys. Rev. Lett.*, vol. 109, no. 18, p. 186404, 2012.

[171] K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck, "Efficient description of strongly correlated electrons with mean-field cost," *Phys. Rev. B*, vol. 89, no. 20, p. 201106, 2014.

[172] E. H. Lieb and F. Wu, "The one-dimensional hubbard model: A reminiscence," *Physica A*, vol. 321, no. 1, pp. 1–27, 2003.

[173] O. Gunnarsson and K. Schönhammer, "Density-functional treatment of an exactly solvable semiconductor model," *Phys. Rev. Lett.*, vol. 56, no. 18, p. 1968, 1986.

[174] A. Schindlmayr and R. Godby, "Density-functional theory and the v-representability problem for model strongly correlated electron systems," *Phys. Rev. B*, vol. 51, no. 16, p. 10427, 1995.

[175] R. López-Sandoval and G. Pastor, "Density-matrix functional theory of the hubbard model: An exact numerical study," *Phys. Rev. B*, vol. 61, no. 3, p. 1764, 2000.

[176] S. F. Boys, "Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another," *Rev. Mod. Phys.*, vol. 32, no. 2, p. 296, 1960.

[177] P. C. Hansen, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, vol. 4. SIAM, 1998.

[178] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.*, vol. 69, no. 19, p. 2863, 1992.

[179] G. K.-L. Chan and M. Head-Gordon, "Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group," *J. Chem. Phys.*, vol. 116, no. 11, pp. 4462–4476, 2002.

[180] C.-C. Chang and S. Zhang, "Spatially inhomogeneous phase in the two-dimensional repulsive hubbard model," *Phys. Rev. B*, vol. 78, no. 16, p. 165101, 2008.

[181] K. Ohno, "Some remarks on the pariser-parr-pople method," *Theor. Chim. Acta*, vol. 2, no. 3, pp. 219–227, 1964.

[182] Y. N. Kalugina, V. N. Cherepanov, M. A. Buldakov, N. Zvereva-Loëte, and V. Boudon, "Theoretical investigation of the ethylene dimer: interaction energy and dipole moment," *J. Comput. Chem.*, vol. 33, no. 3, pp. 319–330, 2012.

[183] N. L. Allinger, Y. H. Yuh, and J. H. Lii, "Molecular mechanics. the mm3 force field for hydrocarbons. 1," *J. Am. Chem. Soc.*, vol. 111, no. 23, pp. 8551–8566, 1989.

[184] J. F. Dobson, "Beyond pairwise additivity in london dispersion interactions," *Int. J. Quantum Chem.*, vol. 114, no. 18, pp. 1157–1161, 2014.

[185] A. M. Reilly and A. Tkatchenko, "van der waals dispersion interactions in molecular materials: beyond pairwise additivity," *Chem. Sci.*, vol. 6, no. 6, pp. 3289–3301, 2015.

[186] T. Zhu, P. de Silva, and T. Van Voorhis, "Self-attractive hartree decomposition: Partitioning electron density into smooth localized fragments," *J. Chem. Theory Comput.*, 2017.

[187] B. Paulus, "The method of incrementsâĂŤa wavefunction-based ab initio correlation method for solids," *Phys. Rep.*, vol. 428, no. 1, pp. 1–52, 2006.

[188] L. Pauling, *The nature of the chemical bond.* Cornell university press: Ithaca, NY, 3 ed., 1960.

[189] C. A. Coulson, *Valence.* Oxford University Press, 1962.

[190] R. G. Parr, P. W. Ayers, and R. F. Nalewajski, "What is an atom in a molecule?," *J. Phys. Chem. A*, vol. 109, no. 17, pp. 3957–3959, 2005.

[191] J. Thiele, "Zur kenntniss der ungesättigten verbindungen. theorie der ungesättigten und aromatischen verbindungen," *Eur. J. Org. Chem.*, vol. 306, no. 1-2, pp. 87–142, 1899.

[192] R. S. Mulliken, "Intensities of electronic transitions in molecular spectra iv. cyclic dienes and hyperconjugation," *J. Chem. Phys.*, vol. 7, no. 5, pp. 339–352, 1939.

[193] M. Jansen and U. Wedig, "A piece of the picture–misunderstanding of chemical concepts," *Angew. Chem. Int. Ed.*, vol. 47, no. 52, pp. 10026–10029, 2008.

[194] J. F. Gonthier, S. N. Steinmann, M. D. Wodrich, and C. Corminboeuf, "Quantification of "fuzzy" chemical concepts: a computational perspective," *Chem. Soc. Rev.*, vol. 41, no. 13, pp. 4671–4687, 2012.

[195] B. B. Averkiev, D. Y. Zubarev, L.-M. Wang, W. Huang, L.-S. Wang, and A. I. Boldyrev, "Carbon avoids hypercoordination in cb6-, cb62-, and c2b5- planar carbon-boron clusters," *J. Am. Chem. Soc.*, vol. 130, no. 29, pp. 9248–9250, 2008.

[196] A. E. Reed, L. A. Curtiss, and F. Weinhold, "Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint," *Chem. Rev.*, vol. 88, no. 6, pp. 899–926, 1988.

[197] P. R. Horn, Y. Mao, and M. Head-Gordon, "Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals," *Phys. Chem. Chem. Phys.*, vol. 18, no. 33, pp. 23067–23079, 2016.

[198] F. De Proft, C. Van Alsenoy, A. Peeters, W. Langenaeker, and P. Geerlings, "Atomic charges, dipole moments, and fukui functions using the hirshfeld partitioning of the electron density," *J. Comput. Chem.*, vol. 23, no. 12, pp. 1198–1209, 2002.

[199] C. Edmiston and K. Ruedenberg, "Localized atomic and molecular orbitals," *Rev. Mod. Phys.*, vol. 35, no. 3, p. 457, 1963.

[200] J. Pipek and P. G. Mezey, "A fast intrinsic localization procedure applicable for abinitio and semiempirical linear combination of atomic orbital wave functions," *J. Chem. Phys.*, vol. 90, no. 9, pp. 4916–4926, 1989.

[201] I.-M. Høyvik, B. Jansik, and P. Jørgensen, "Orbital localization using fourth central moment minimization," *J. Chem. Phys.*, vol. 137, no. 22, p. 224114, 2012.

[202] S. Liu, J. M. Perez-Jorda, and W. Yang, "Nonorthogonal localized molecular orbitals in electronic structure theory," *J. Chem. Phys.*, vol. 112, no. 4, pp. 1634–1644, 2000.

[203] A. E. Reed and F. Weinhold, "Natural localized molecular orbitals," *J. Chem. Phys.*, vol. 83, no. 4, pp. 1736–1740, 1985.

[204] P. de Silva, M. Giebułtowski, and J. Korchowiec, "Fast orbital localization scheme in molecular fragments resolution," *Phys. Chem. Chem. Phys.*, vol. 14, no. 2, pp. 546–552, 2012.

[205] A. Imamura, Y. Aoki, and K. Maekawa, "A theoretical synthesis of polymers by using uniform localization of molecular orbitals: proposal of an elongation method," *J. Chem. Phys.*, vol. 95, no. 7, pp. 5419–5431, 1991.

[206] M. Schütz, G. Hetzer, and H.-J. Werner, "Low-order scaling local electron correlation methods. i. linear scaling local mp2," *J. Chem. Phys.*, vol. 111, no. 13, pp. 5691–5705, 1999.

[207] M. Schütz and H.-J. Werner, "Low-order scaling local electron correlation methods. iv. linear scaling local coupled-cluster (lccsd)," *J. Chem. Phys.*, vol. 114, no. 2, pp. 661–681, 2001.

[208] R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski, *Linear-scaling techniques in computational chemistry and physics: Methods and applications.* Springer Science+ Business Media BV, 2011.

[209] T. A. Wesolowski and Y. A. Wang, *Recent Progress in Orbital-Free Density Functional Theory*, vol. 6. World Scientific, 2013.

[210] T. Helgaker, H. Larsen, J. Olsen, and P. Jørgensen, "Direct optimization of the ao density matrix in hartree–fock and kohn–sham theories," *Chem. Phys. Lett.*, vol. 327, no. 5, pp. 397–403, 2000.

[211] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, no. 3B, p. B864, 1964.

[212] R. F. Bader, *Atoms in molecules.* Wiley Online Library, 1990.

[213] R. F. Bader, "A quantum theory of molecular structure and its applications," *Chem. Rev.*, vol. 91, no. 5, pp. 893–928, 1991.

[214] H. J. Bohórquez and R. J. Boyd, "A localized electrons detector for atomic and molecular systems," *Theor. Chem. Acc.*, vol. 127, no. 4, pp. 393–400, 2010.

[215] P. de Silva, J. Korchowiec, and T. A. Wesolowski, "Revealing the bonding pattern from the molecular electron density using single exponential decay detector: An orbital-free alternative to the electron localization function," *ChemPhysChem*, vol. 13, no. 15, pp. 3462–3465, 2012.

[216] P. de Silva, J. Korchowiec, J. Ram, and T. A. Wesolowski, "Extracting information about chemical bonding from molecular electron densities via single exponential decay detector (sedd)," *CHIMIA*, vol. 67, no. 4, pp. 253–256, 2013.

[217] P. de Silva, J. Korchowiec, and T. A. Wesolowski, "Atomic shell structure from the single-exponential decay detector," *J. Chem. Phys.*, vol. 140, no. 16, p. 164301, 2014.

[218] P. de Silva and C. Corminboeuf, "Simultaneous visualization of covalent and noncovalent interactions using regions of density overlap," *J. Chem. Theory Comput.*, vol. 10, no. 9, p. 3745, 2014.

[219] S. S. Shaik and P. C. Hiberty, *A chemist's guide to valence bond theory.* John Wiley & Sons, 2007.

[220] V. A. Miransky, *Dynamical symmetry breaking in quantum field theories.* World Scientific, 1994.

[221] J. P. Perdew, A. Ruzsinszky, J. Sun, and M. R. Pederson, "Chapter one-paradox of self-interaction correction: How can anything so right be so wrong?," *Adv. At. Mol. Opt. Phys.*, vol. 64, pp. 1–14, 2015.

[222] S. Klüpfel, P. Klüpfel, and H. Jónsson, "Importance of complex orbitals in calculating the self-interaction-corrected ground state of atoms," *Phys. Rev. A*, vol. 84, no. 5, p. 050501, 2011.

[223] S. Klüpfel, P. Klüpfel, and H. Jónsson, "The effect of the perdew-zunger self-interaction correction to density functionals on the energetics of small molecules," *J. Chem. Phys.*, vol. 137, no. 12, p. 124102, 2012.

[224] S. Lehtola and H. Jónsson, "Variational, self-consistent implementation of the perdew–zunger self-interaction correction with complex optimal orbitals," *J. Chem. Theory Comput.*, vol. 10, no. 12, pp. 5324–5337, 2014.

[225] C. R. Jacob and J. Neugebauer, "Subsystem density-functional theory," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 4, no. 4, pp. 325–362, 2014.

[226] J. L. Whitten, "Coulombic potential energy integrals and approximations," *J. Chem. Phys.*, vol. 58, no. 10, pp. 4496–4501, 1973.

[227] F. Weigend, A. Köhn, and C. Hättig, "Efficient use of the correlation consistent basis sets in resolution of the identity mp2 calculations," *J. Chem. Phys.*, vol. 116, no. 8, pp. 3175–3183, 2002.

[228] B. I. Dunlap, J. W. Connolly, and J. R. Sabin, "On the applicability of lcao-xα methods to molecules containing transition metal atoms: The nickel atom and nickel hydride," *Int. J. Quantum Chem.*, vol. 12, no. S11, pp. 81–87, 1977.

[229] B. I. Dunlap, J. Connolly, and J. Sabin, "On some approximations in applications of xα theory," *J. Chem. Phys.*, vol. 71, no. 8, pp. 3396–3402, 1979.

[230] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kús, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr., H. Dop, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, P. A. Pieniazek, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, N. Sergueev, S. M. Sharada, S. Sharmaa, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White,

C. F. Williams, V. Vanovschi, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhou, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xua, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the q-chem 4 program package," *Mol. Phys.*, vol. 113, pp. 184–215, 2015.

[231] T. H. Dunning Jr, "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen," *J. Chem. Phys.*, vol. 90, no. 2, pp. 1007–1023, 1989.

[232] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions," *J. Chem. Phys.*, vol. 96, no. 9, pp. 6796–6806, 1992.

[233] V. N. Staroverov, G. E. Scuseria, and E. R. Davidson, "Optimized effective potentials yielding hartree–fock energies and densities," *J. Chem. Phys.*, vol. 124, no. 14, p. 141103, 2006.

[234] S. Hirata, S. Ivanov, I. Grabowski, R. J. Bartlett, K. Burke, and J. D. Talman, "Can optimized effective potentials be determined uniquely?," *J. Chem. Phys.*, vol. 115, no. 4, pp. 1635–1649, 2001.

[235] A. Heßelmann, A. W. Götz, F. Della Sala, and A. Görling, "Numerically stable optimized effective potential method with balanced gaussian basis sets," *J. Chem. Phys.*, vol. 127, no. 5, p. 054102, 2007.

[236] F. A. Bulat, T. Heaton-Burgess, A. J. Cohen, and W. Yang, "Optimized effective potentials from electron densities in finite basis sets," *J. Chem. Phys.*, vol. 127, no. 17, p. 174101, 2007.

[237] C. R. Jacob, "Unambiguous optimization of effective potentials in finite basis sets," *J. Chem. Phys.*, vol. 135, no. 24, p. 244102, 2011.

[238] W. Humphrey, A. Dalke, and K. Schulten, "Vmd: visual molecular dynamics," *J. Mol. Graph.*, vol. 14, no. 1, pp. 33–38, 1996.

[239] G. Schaftenaar and J. H. Noordik, "Molden: a pre-and post-processing program for molecular and electronic structures," *J. Comput. Aid. Mol. Des.*, vol. 14, no. 2, pp. 123–134, 2000.

[240] P. de Silva and T. A. Wesolowski, "Pure-state noninteracting v-representability of electron densities from kohn-sham calculations with finite basis sets," *Phys. Rev. A*, vol. 85, no. 3, p. 032518, 2012.

[241] S. Lehtola and H. Jónsson, "Pipek–mezey orbital localization using various partial charge estimates," *J. Chem. Theory Comput.*, vol. 10, no. 2, pp. 642–649, 2014.

[242] A. D. Clauss, S. F. Nelsen, M. Ayoub, J. W. Moore, C. R. Landis, and F. Weinhold, "Rabbit-ears hybrids, vsepr sterics, and other orbital anachronisms," *Chem. Educ. Res. Pract.*, vol. 15, no. 4, pp. 417–434, 2014.

[243] P. C. Hiberty, D. Danovich, and S. Shaik, "Comment on "rabbit-ears hybrids, vsepr sterics, and other orbital anachronisms". a reply to a criticism," *Chem. Educ. Res. Pract.*, vol. 16, no. 3, pp. 689–693, 2015.

[244] L. Pauling, "The nature of the chemical bond. application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules," *J. Am. Chem. Soc.*, vol. 53, no. 4, pp. 1367–1400, 1931.

[245] E. R. Davidson and S. Chakravorty, "A test of the hirshfeld definition of atomic charges and moments," *Theor. Chem. Acc.*, vol. 83, no. 5, pp. 319–330, 1992.

[246] S. Tsuzuki and H. P. Lüthi, "Interaction energies of van der waals and hydrogen bonded systems calculated using density functional theory: Assessing the pw91 model," *J. Chem. Phys.*, vol. 114, no. 9, pp. 3949–3957, 2001.

[247] A. D. Boese, J. M. Martin, and W. Klopper, "Basis set limit coupled cluster study of h-bonded systems and assessment of more approximate methods," *J. Phys. Chem. A*, vol. 111, no. 43, pp. 11122–11133, 2007.

[248] O. A. Vydrov and T. Van Voorhis, "Benchmark assessment of the accuracy of several van der waals density functionals," *J. Chem. Theory Comput.*, vol. 8, no. 6, pp. 1929–1934, 2012.

[249] S. J. Grabowski, "What is the covalency of hydrogen bonding?," *Chem. Rev.*, vol. 111, no. 4, pp. 2597–2625, 2011.

[250] S. Gómez, J. Nafziger, A. Restrepo, and A. Wasserman, "Partition-dft on the water dimer," *J. Chem. Phys.*, vol. 146, no. 7, p. 074106, 2017.

[251] R. Parthasarathi, V. Subramanian, and N. Sathyamurthy, "Hydrogen bonding without borders: an atoms-in-molecules perspective," *J. Phys. Chem. A*, vol. 110, no. 10, pp. 3349–3351, 2006.

[252] J. R. Lane, A. S. Hansen, K. Mackeprang, and H. G. Kjaergaard, "Kinetic energy density as a predictor of hydrogen-bonded oh-stretching frequencies," *J. Phys. Chem. A*, vol. 121, no. 18, pp. 3452–3460, 2017.

[253] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legnon, B. Mennucci, and D. J. Nesbitt, "Definition of the hydrogen bond (iupac recommendations 2011)," *Pure Appl. Chem.*, vol. 83, no. 8, pp. 1637–1641, 2011.

[254] S. F. Boys and F. d. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors," *Mol. Phys.*, vol. 19, no. 4, pp. 553–566, 1970.

[255] K. Morokuma, "Why do molecules interact? the origin of electron donor-acceptor complexes, hydrogen bonding and proton affinity," *Acc. Chem. Res.*, vol. 10, no. 8, pp. 294–300, 1977.

[256] S. J. Grabowski, *Hydrogen bonding: new insights*, vol. 3. Springer, 2006.

[257] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen, and W. Yang, "Revealing noncovalent interactions," *J. Am. Chem. Soc.*, vol. 132, no. 18, pp. 6498–6506, 2010.

[258] T. Lu and F. Chen, "Multiwfn: a multifunctional wavefunction analyzer," *J. Comput. Chem.*, vol. 33, no. 5, pp. 580–592, 2012.

[259] G. Saleh, L. Lo Presti, C. Gatti, and D. Ceresoli, "Ncimilano: an electron-density-based code for the study of noncovalent interactions," *J. Appl. Crystallogr.*, vol. 46, no. 5, pp. 1513–1517, 2013.

[260] J. R. Lane, J. Contreras-García, J.-P. Piquemal, B. J. Miller, and H. G. Kjaergaard, "Are bond critical points really critical for hydrogen bonding?," *J. Chem. Theory Comput.*, vol. 9, no. 8, pp. 3263–3266, 2013.

[261] D. L. Howard, P. Jørgensen, and H. G. Kjaergaard, "Weak intramolecular interactions in ethylene glycol identified by vapor phase oh- stretching overtone spectroscopy," *J. Am. Chem. Soc.*, vol. 127, no. 48, pp. 17096–17103, 2005.

[262] D. L. Howard and H. G. Kjaergaard, "Influence of intramolecular hydrogen bond strength on oh-stretching overtones," *J. Phys. Chem. A*, vol. 110, no. 34, pp. 10245–10250, 2006.

[263] J. Gregory, D. Clary, K. Liu, M. Brown, and R. Saykally, "The water dipole moment in water clusters," *Science*, vol. 275, no. 5301, pp. 814–817, 1997.

[264] G. D. Purvis III and R. J. Bartlett, "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples," *J. Chem. Phys.*, vol. 76, no. 4, pp. 1910–1918, 1982.

[265] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, "Pyscf: the python-based simulations of chemistry framework," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 8, no. 1, 2018.

[266] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. K.-L. Chan, "The ab-initio density matrix renormalization group in practice," *J. Chem. Phys.*, vol. 142, no. 3, p. 034102, 2015.

[267] X. Zheng, M. Liu, E. R. Johnson, J. Contreras-García, and W. Yang, "Delocalization error of density-functional approximations: A distinct manifestation in hydrogen molecular chains," *J. Chem. Phys.*, vol. 137, no. 21, p. 214106, 2012.

[268] N. Q. Su and X. Xu, "Error accumulations in adhesive energies of dihydrogen molecular chains: performances of the xyg3 type of doubly hybrid density functionals," *J. Phys. Chem. A*, vol. 119, no. 9, pp. 1590–1599, 2014.

[269] M.-C. Kim, E. Sim, and K. Burke, "Understanding and reducing errors in density functional calculations," *Phys. Rev. Lett.*, vol. 111, no. 7, p. 073003, 2013.

[270] G. Knizia and G. K.-L. Chan, "Density matrix embedding: A strong-coupling quantum embedding theory," *J. Chem. Theory. Comput.*, vol. 9, no. 3, pp. 1428–1432, 2013.

[271] I. W. Bulik, T. M. Henderson, and G. E. Scuseria, "Can single-reference coupled cluster theory describe static correlation?," *J. Chem. Theory. Comput.*, vol. 11, no. 7, pp. 3171–3179, 2015.

[272] T. Zhu and T. Van Voorhis, "Charge recombination in phosphorescent organic light-emitting diode host–guest systems through qm/mm simulations," *J. Phys. Chem. C*, vol. 120, no. 36, pp. 19987–19994, 2016.

[273] M. A. Baldo, D. O'brien, Y. You, A. Shoustikov, S. Sibley, M. Thompson, and S. Forrest, "Highly efficient phosphorescent emission from organic electroluminescent devices," *Nature*, vol. 395, no. 6698, pp. 151–154, 1998.

[274] M. Baldo, S. Lamansky, P. Burrows, M. Thompson, and S. Forrest, "Very high-efficiency green organic light-emitting devices based on electrophosphorescence," *Appl. Phys. Lett.*, vol. 75, p. 4, 1999.

[275] C. Adachi, M. A. Baldo, S. R. Forrest, and M. E. Thompson, "High-efficiency organic electrophosphorescent devices with tris (2-phenylpyridine) iridium doped into electron-transporting materials," *Appl. Phys. Lett.*, vol. 77, no. 6, p. 904, 2000.

[276] Y. Sun, N. C. Giebink, H. Kanno, B. Ma, M. E. Thompson, and S. R. Forrest, "Management of singlet and triplet excitons for efficient white organic light-emitting devices," *Nature*, vol. 440, no. 7086, pp. 908–912, 2006.

[277] S. Reineke, F. Lindner, G. Schwartz, N. Seidler, K. Walzer, B. Lüssem, and K. Leo, "White organic light-emitting diodes with fluorescent tube efficiency," *Nature*, vol. 459, no. 7244, pp. 234–238, 2009.

[278] A. Brown, K. Pichler, N. Greenham, D. Bradley, R. H. Friend, and A. Holmes, "Optical spectroscopy of triplet excitons and charged excitations in poly (p-phenylenevinylene) light-emitting diodes," *Chem. Phys. Lett.*, vol. 210, no. 1, pp. 61–66, 1993.

[279] M. Baldo, D. O'brien, M. Thompson, and S. Forrest, "Excitonic singlet-triplet ratio in a semiconducting organic thin film," *Phys. Rev. B*, vol. 60, no. 20, p. 14422, 1999.

[280] V. Cleave, G. Yahioglu, P. L. Barny, R. H. Friend, and N. Tessler, "Harvesting singlet and triplet energy in polymer leds," *Adv. Mater.*, vol. 11, no. 4, pp. 285–288, 1999.

[281] H. Yersin, "Triplet emitters for oled applications. mechanisms of exciton trapping and control of emission properties," in *Transition Metal and Rare Earth Compounds*, pp. 1–26, Springer, 2004.

[282] C. Adachi, M. A. Baldo, M. E. Thompson, and S. R. Forrest, "Nearly 100% internal phosphorescence efficiency in an organic light-emitting device," *J. Appl. Phys*, vol. 90, pp. 5048–5051, 2001.

[283] E. L. Williams, K. Haavisto, J. Li, and G. E. Jabbour, "Excimer-based white phosphorescent organic light-emitting diodes with nearly 100% internal quantum efficiency," *Adv. Mater.*, vol. 19, no. 2, pp. 197–202, 2007.

[284] L. Xiao, S.-J. Su, Y. Agata, H. Lan, and J. Kido, "Nearly 100% internal quantum efficiency in an organic blue-light electrophosphorescent device using a weak electron transporting material with a wide energy gap," *Adv. Mater.*, vol. 21, no. 12, pp. 1271–1274, 2009.

[285] D. Tanaka, H. Sasabe, Y.-J. Li, S.-J. Su, T. Takeda, and J. Kido, "Ultra high efficiency green organic light-emitting devices," *Jpn. J. Appl. Phys.*, vol. 46, no. 1L, p. L10, 2007.

[286] S.-J. Su, C. Cai, and J. Kido, "Rgb phosphorescent organic light-emitting diodes by using host materials with heterocyclic cores: effect of nitrogen atom orientations," *Chem. Mater.*, vol. 23, no. 2, pp. 274–284, 2010.

[287] S. O. Jeon, S. E. Jang, H. S. Son, and J. Y. Lee, "External quantum efficiency above 20% in deep blue phosphorescent organic light-emitting diodes," *Adv. Mater.*, vol. 23, no. 12, pp. 1436–1441, 2011.

[288] Y. Tao, Q. Wang, L. Ao, C. Zhong, J. Qin, C. Yang, and D. Ma, "Molecular design of host materials based on triphenylamine/oxadiazole hybrids for excellent deep-red phosphorescent organic light-emitting diodes," *J. Mater. Chem.*, vol. 20, no. 9, pp. 1759–1765, 2010.

[289] H. Uoyama, K. Goushi, K. Shizu, H. Nomura, and C. Adachi, "Highly efficient organic light-emitting diodes from delayed fluorescence," *Nature*, vol. 492, no. 7428, pp. 234–238, 2012.

[290] H. Kaji, H. Suzuki, T. Fukushima, K. Shizu, K. Suzuki, S. Kubo, T. Komino, H. Oiwa, F. Suzuki, A. Wakamiya, Y. Murata, and C. Adachi, "Purely organic electroluminescent material realizing 100% conversion from electricity to light," *Nat. Commun.*, vol. 6, 2015.

[291] Y. Tao, C. Yang, and J. Qin, "Organic host materials for phosphorescent organic light-emitting diodes," *Chem. Soc. Rev.*, vol. 40, no. 5, pp. 2943–2970, 2011.

[292] A. Chaskar, H.-F. Chen, and K.-T. Wong, "Bipolar host materials: a chemical approach for highly efficient electrophosphorescent devices," *Adv. Mater.*, vol. 23, no. 34, pp. 3876–3895, 2011.

[293] S.-J. Yeh, M.-F. Wu, C.-T. Chen, Y.-H. Song, Y. Chi, M.-H. Ho, S.-F. Hsu, and C. H. Chen, "New dopant and host materials for blue-light-emitting phosphorescent organic electroluminescent devices," *Adv. Mater.*, vol. 17, no. 3, pp. 285–+, 2005.

[294] S. Tokito, T. Iijima, Y. Suzuri, H. Kita, T. Tsuzuki, and F. Sato, "Confinement of triplet energy on phosphorescent molecules for highly-efficient organic blue-light-emitting devices," *Appl. Phys. Lett.*, vol. 83, no. 3, pp. 569–571, 2003.

[295] H. Nakanotani, K. Masui, J. Nishide, T. Shibata, and C. Adachi, "Promising operational stability of high-efficiency organic light-emitting diodes based on thermally activated delayed fluorescence," *Sci. Rep.*, vol. 3, 2013.

[296] W. Li, J. Li, F. Wang, Z. Gao, and S. Zhang, "Universal host materials for high-efficiency phosphorescent and delayed-fluorescence oleds," *ACS Appl. Mater. Interfaces*, vol. 7, no. 47, pp. 26206–26216, 2015.

[297] M. Einzinger, T. Zhu, P. de Silva, C. Belger, T. M. Swager, T. Van Voorhis, and M. A. Baldo, "Shorter exciton lifetimes via an external heavy-atom effect: Alleviating the effects of bimolecular processes in organic light-emitting diodes," *Adv. Mater.*, vol. 29, no. 40, 2017.

[298] T. Förster, "Zwischenmolekulare energiewanderung und fluoreszenz," *Ann. Phys. (Berlin)*, vol. 437, no. 1-2, pp. 55–75, 1948.

[299] D. L. Dexter, "A theory of sensitized luminescence in solids," *J. Chem. Phys.*, vol. 21, no. 5, pp. 836–850, 1953.

[300] M. Baldo, M. Thompson, and S. Forrest, "High-efficiency fluorescent organic light-emitting devices using a phosphorescent sensitizer," *Nature*, vol. 403, no. 6771, pp. 750–753, 2000.

[301] Y. Seino, H. Sasabe, Y.-J. Pu, and J. Kido, "High-performance blue phosphorescent oleds using energy transfer from exciplex," *Adv. Mater.*, vol. 26, no. 10, pp. 1612–1616, 2014.

[302] H. Nakanotani, T. Higuchi, T. Furukawa, K. Masui, K. Morimoto, M. Numata, H. Tanaka, Y. Sagara, T. Yasuda, and C. Adachi, "High-efficiency organic light-emitting diodes with fluorescent emitters," *Nat. Commun.*, vol. 5, 2014.

[303] X. Gong, J. C. Ostrowski, D. Moses, G. Bazan, and A. Heeger, "Electrophosphorescence from a polymer guest–host system with an iridium complex as guest: Förster energy transfer and charge trapping," *Adv. Funct. Mater.*, vol. 13, no. 6, pp. 439–444, 2003.

[304] R. Holmes, B. D'Andrade, S. Forrest, X. Ren, J. Li, and M. Thompson, "Efficient, deep-blue organic electrophosphorescence by guest charge trapping," *Appl. Phys. Lett.*, vol. 83, no. 18, pp. 3818–3820, 2003.

[305] W. S. Jeon, T. J. Park, S. Y. Kim, R. Pode, J. Jang, and J. H. Kwon, "Ideal host and guest system in phosphorescent oleds," *Org. Electron.*, vol. 10, no. 2, pp. 240–246, 2009.

[306] B. Diouf, W. S. Jeon, R. Pode, and J. H. Kwon, "Efficiency control in iridium complex-based phosphorescent light-emitting diodes," *Adv. Mater. Sci. Eng.*, vol. 2012, 2012.

[307] Y. J. Cho, K. S. Yook, and J. Y. Lee, "A universal host material for high external quantum efficiency close to 25% and long lifetime in green fluorescent and phosphorescent oleds," *Adv. Mater.*, vol. 26, no. 24, pp. 4050–4055, 2014.

[308] H.-H. Chou and C.-H. Cheng, "A highly efficient universal bipolar host for blue, green, and red phosphorescent oleds," *Adv. Mater.*, vol. 22, no. 22, pp. 2468–2471, 2010.

[309] F. May, M. Al-Helwi, B. Baumeier, W. Kowalsky, E. Fuchs, C. Lennartz, and D. Andrienko, "Design rules for charge-transport efficient host materials for phosphorescent organic light-emitting diodes," *J. Am. Chem. Soc.*, vol. 134, no. 33, pp. 13818–13822, 2012.

[310] F. May, B. Baumeier, C. Lennartz, and D. Andrienko, "Can lattice models predict the density of states of amorphous organic semiconductors?," *Phys. Rev. Lett.*, vol. 109, no. 13, p. 136401, 2012.

[311] P. Kordt, J. J. van der Holst, M. Al Helwi, W. Kowalsky, F. May, A. Badinski, C. Lennartz, and D. Andrienko, "Modeling of organic light emitting diodes: From molecular to device properties," *Adv. Funct. Mater.*, vol. 25, no. 13, pp. 1955–1971, 2015.

[312] K. M. Pelzer and S. B. Darling, "Charge generation in organic photovoltaics: A review of theory and computation," *Mol. Syst. Des. & Eng.*, vol. 1, no. 1, pp. 10–24, 2016.

[313] J.-L. Brédas, J. E. Norton, J. Cornil, and V. Coropceanu, "Molecular understanding of organic solar cells: the challenges," *Acc. Chem. Res.*, vol. 42, no. 11, pp. 1691–1699, 2009.

[314] J. Åqvist and A. Warshel, "Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches," *Chem. Rev.*, vol. 93, no. 7, pp. 2523–2544, 1993.

[315] S. Difley, L.-P. Wang, S. Yeganeh, S. R. Yost, and T. Van Voorhis, "Electronic properties of disordered organic semiconductors via qm/mm simulations," *Acc. Chem. Res.*, vol. 43, no. 7, pp. 995–1004, 2010.

[316] S. R. Yost, L.-P. Wang, and T. Van Voorhis, "Molecular insight into the energy levels at the organic donor/acceptor interface: A quantum mechanics/molecular mechanics study," *J. Phys. Chem. C*, vol. 115, no. 29, pp. 14431–14436, 2011.

[317] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, "Packmol: a package for building initial configurations for molecular dynamics simulations," *J. Comput. Chem.*, vol. 30, no. 13, pp. 2157–2164, 2009.

[318] H. J. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. Haak, "Molecular dynamics with coupling to an external bath," *J. Chem. Phys.*, vol. 81, no. 8, pp. 3684–3690, 1984.

[319] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh ewald method," *J. Chem. Phys.*, vol. 103, no. 19, pp. 8577–8593, 1995.

[320] B. Hess, C. Kutzner, D. Van Der Spoel, and E. Lindahl, "Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation," *J. Chem. Theory Comput.*, vol. 4, no. 3, pp. 435–447, 2008.

[321] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids," *J. Am. Chem. Soc.*, vol. 118, no. 45, pp. 11225–11236, 1996.

[322] C. M. Breneman and K. B. Wiberg, "Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis," *J. Comput. Chem.*, vol. 11, no. 3, pp. 361–373, 1990.

[323] P. C. Hariharan and J. A. Pople, "The influence of polarization functions on molecular orbital hydrogenation energies," *Theor. Chem. Acc.*, vol. 28, no. 3, pp. 213–222, 1973.

[324] P. J. Hay and W. R. Wadt, "Ab initio effective core potentials for molecular calculations. potentials for k to au including the outermost core orbitals," *J. Chem. Phys.*, vol. 82, no. 1, pp. 299–310, 1985.

[325] J. M. Herbert, L. D. Jacobson, K. U. Lao, and M. A. Rohrdanz, "Rapid computation of intermolecular interactions in molecular and ionic clusters: Self-consistent polarization plus symmetry-adapted perturbation theory," *Phys. Chem./Chem. Phys.*, vol. 14, no. 21, pp. 7679–7699, 2012.

[326] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "Charmm: A program for macromolecular energy, minimization, and dynamics calculations," *J. Comput. Chem.*, vol. 4, no. 2, pp. 187–217, 1983.

[327] H. L. Woodcock, M. Hodošček, A. T. Gilbert, P. M. Gill, H. F. Schaefer, and B. R. Brooks, "Interfacing q-chem and charmm to perform qm/mm reaction path calculations," *J. Comput. Chem.*, vol. 28, no. 9, pp. 1485–1502, 2007.

[328] C. Adamo, G. E. Scuseria, and V. Barone, "Accurate excitation energies from time-dependent density functional theory: Assessing the pbe0 model," *J. Chem. Phys.*, vol. 111, no. 7, pp. 2889–2899, 1999.

[329] I. Hill, A. Kahn, Z. Soos, and R. Pascal Jr, "Charge-separation energy in films of π-conjugated organic molecules," *Chem. Phys. Lett.*, vol. 327, no. 3, pp. 181–188, 2000.

[330] S. Krause, M. Casu, A. Schöll, and E. Umbach, "Determination of transport levels of organic semiconductors by ups and ips," *New J. Phys.*, vol. 10, no. 8, p. 085001, 2008.

[331] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Stat.*, vol. 27, no. 3, pp. 832–837, 1956.

[332] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Stat.*, vol. 33, no. 3, pp. 1065–1076, 1962.

[333] P. de Silva and T. Van Voorhis, "Qm/mm study of static and dynamic energetic disorder in the emission layer of an organic light-emitting diode," *J. Phys. Chem. Lett.*, vol. 9, pp. 1329–1334, 2018.

[334] B. W. D'Andrade, S. Datta, S. R. Forrest, P. Djurovich, E. Polikarpov, and M. E. Thompson, "Relationship between the ionization and oxidation potentials of molecular organic semiconductors," *Org. Electron.*, vol. 6, no. 1, pp. 11–20, 2005.

[335] H. Yoshida and K. Yoshizaki, "Electron affinities of organic materials used for organic light-emitting diodes: A low-energy inverse photoemission study," *Org. Electron.*, vol. 20, pp. 24–30, 2015.

[336] R. Ross, J. Powers, T. Atashroo, W. Ermler, L. LaJohn, and P. Christiansen, "Ab initio relativistic effective potentials with spin–orbit operators. iv. cs through rn," *J. Chem. Phys.*, vol. 93, no. 9, pp. 6654–6670, 1990.

[337] S. Scholz, D. Kondakov, B. LuÌĺssem, and K. Leo, "Degradation mechanisms and reactions in organic light-emitting devices," *Chem. Rev.*, vol. 115, no. 16, pp. 8449–8503, 2015.

[338] D. Hait, T. Zhu, D. P. McMahon, and T. Van Voorhis, "Prediction of excited-state energies and singlet–triplet gaps of charge-transfer states using a restricted open-shell kohn–sham approach," *J. Chem. Theory Comput.*, vol. 12, no. 7, pp. 3353–3359, 2016.

[339] J. Fine, K. Diri, A. Krylov, C. Nemirow, Z. Lu, and C. Wittig, "Electronic structure of tris (2-phenylpyridine) iridium: electronically excited and ionized states," *Mol. Phys.*, vol. 110, no. 15-16, pp. 1849–1862, 2012.

[340] Y. Kawamura, J. Brooks, J. J. Brown, H. Sasabe, and C. Adachi, "Intermolecular interaction and a concentration-quenching mechanism of phosphorescent ir (iii) complexes in a solid film," *Phys. Rev. Lett.*, vol. 96, no. 1, p. 017404, 2006.

[341] M. Cossi, N. Rega, G. Scalmani, and V. Barone, "Energies, structures, and electronic properties of molecules in solution with the c-pcm solvation model," *J. Comput. Chem.*, vol. 24, no. 6, pp. 669–681, 2003.

[342] M. A. Baldo, C. Adachi, and S. R. Forrest, "Transient analysis of organic electrophosphorescence. ii. transient analysis of triplet-triplet annihilation," *Phys. Rev. B*, vol. 62, no. 16, p. 10967, 2000.

[343] K. Sato, K. Shizu, K. Yoshimura, A. Kawada, H. Miyazaki, and C. Adachi, "Organic luminescent molecule with energetically equivalent singlet and triplet excited states for organic light-emitting diodes," *Phys. Rev. Lett.*, vol. 110, no. 24, p. 247401, 2013.

[344] S. Hirata, Y. Sakai, K. Masui, H. Tanaka, S. Y. Lee, H. Nomura, N. Nakamura, M. Yasumatsu, H. Nakanotani, Q. Zhang, K. Shizu, H. Miyazaki, and C. Adachi, "Highly efficient blue electroluminescence based on thermally activated delayed fluorescence," *Nat. Mater.*, vol. 14, no. 3, pp. 330–336, 2015.

[345] Y. Im, M. Kim, Y. J. Cho, J.-A. Seo, K. S. Yook, and J. Y. Lee, "Molecular design strategy of organic thermally activated delayed fluorescence emitters," *Chem. Mater.*, vol. 29, no. 5, pp. 1946–1963, 2017.

[346] Q. Zhang, B. Li, S. Huang, H. Nomura, H. Tanaka, and C. Adachi, "Efficient blue organic light-emitting diodes employing thermally activated delayed fluorescence," *Nat. Photonics*, vol. 8, no. 4, pp. 326–332, 2014.

[347] P. Rajamalli, N. Senthilkumar, P. Gandeepan, P.-Y. Huang, M.-J. Huang, C.-Z. Ren-Wu, C.-Y. Yang, M.-J. Chiu, L.-K. Chu, H.-W. Lin, *et al.*, "A new molecular design based on thermally activated delayed fluorescence for highly efficient organic light emitting diodes," *J. Am. Chem. Soc.*, vol. 138, no. 2, pp. 628–634, 2016.

[348] K. Kawasumi, T. Wu, T. Zhu, H. S. Chae, T. Van Voorhis, M. A. Baldo, and T. M. Swager, "Thermally activated delayed fluorescence materials based on homoconjugation effect of donor–acceptor triptycenes," *J. Am. Chem. Soc.*, vol. 137, no. 37, pp. 11908–11911, 2015.

[349] W. Huang, M. Einzinger, T. Zhu, H. S. Chae, S. Jeon, S.-G. Ihn, M. Sim, S. Kim, M. Su, G. Teverovskiy, *et al.*, "Molecular design of deep blue thermally activated delayed fluorescence materials employing a homoconjugative triptycene scaffold and dihedral angle tuning," *Chem. Mater.*, 2018.

[350] Q. Zhang, J. Li, K. Shizu, S. Huang, S. Hirata, H. Miyazaki, and C. Adachi, "Design of efficient thermally activated delayed fluorescence materials for pure blue organic light emitting diodes," *J. Am. Chem. Soc.*, vol. 134, no. 36, pp. 14706–14709, 2012.

[351] H. Tanaka, K. Shizu, H. Miyazaki, and C. Adachi, "Efficient green thermally activated delayed fluorescence (tadf) from a phenoxazine–triphenyltriazine (pxz–trz) derivative," *Chem. Comm.*, vol. 48, no. 93, pp. 11392–11394, 2012.

[352] A. Dreuw, J. L. Weisman, and M. Head-Gordon, "Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange," *J. Chem. Phys.*, vol. 119, no. 6, pp. 2943–2946, 2003.

[353] A. L. Sobolewski and W. Domcke, "Ab initio study of the excited-state coupled electron–proton-transfer process in the 2-aminopyridine dimer," *Chem. Phys.*, vol. 294, no. 1, pp. 73–83, 2003.

[354] A. Dreuw, G. R. Fleming, and M. Head-Gordon, "Chlorophyll fluorescence quenching by xanthophylls," *Phys. Chem. Chem. Phys.*, vol. 5, no. 15, pp. 3247–3256, 2003.

[355] A. Dreuw and M. Head-Gordon, "Single-reference ab initio methods for the calculation of excited states of large molecules," *Chem. Rev.*, vol. 105, no. 11, pp. 4009–4037, 2005.

[356] L.-S. Cui, H. Nomura, Y. Geng, J. U. Kim, H. Nakanotani, and C. Adachi, "Controlling singlet–triplet energy splitting for deep-blue thermally activated delayed fluorescence emitters," *Angew. Chem. Int. Ed.*, vol. 129, no. 6, pp. 1593–1597, 2017.

[357] N. Harada, H. Uda, K. Nakasuji, and I. Murata, "Interchromophoric homoconjugation effect and intramolecular charge-transfer transition of the triptycene system containing a tetracyanoquinodimethane chromophore," *J. Chem. Soc., Perkin Trans. 2*, no. 10, pp. 1449–1453, 1989.

[358] T. M. Swager, "Iptycenes in the design of high performance polymers," *Acc. Chem. Res.*, vol. 41, no. 9, pp. 1181–1189, 2008.

[359] H.-H. Chou, H.-H. Shih, and C.-H. Cheng, "Triptycene derivatives as high-tg host materials for various electrophosphorescent devices," . *Mater. Chem.*, vol. 20, no. 4, pp. 798–805, 2010.

[360] M. Kim, S. K. Jeon, S.-H. Hwang, and J. Y. Lee, "Stable blue thermally activated delayed fluorescent organic light-emitting diodes with three times longer lifetime than phosphorescent organic light-emitting diodes," *Adv. Mater.*, vol. 27, no. 15, pp. 2515–2520, 2015.

[361] J. W. Sun, J. Y. Baek, K.-H. Kim, C.-K. Moon, J.-H. Lee, S.-K. Kwon, Y.-H. Kim, and J.-J. Kim, "Thermally activated delayed fluorescence from azasiline based intramolecular charge-transfer emitter (dtpdda) and a highly efficient blue light emitting diode," *Chem. Mater.*, vol. 27, no. 19, pp. 6675–6681, 2015.

[362] J. S. Ward, R. S. Nobuyasu, A. S. Batsanov, P. Data, A. P. Monkman, F. B. Dias, and M. R. Bryce, "The interplay of thermally activated delayed fluorescence (tadf) and room temperature organic phosphorescence in sterically-constrained donor–acceptor charge-transfer molecules," *Chem. Comm.*, vol. 52, no. 12, pp. 2612–2615, 2016.

[363] M. Kim, S. K. Jeon, S.-H. Hwang, S.-S. Lee, E. Yu, and J. Y. Lee, "Correlation of molecular structure with photophysical properties and device performances of thermally activated delayed fluorescent emitters," *J. Phys. Chem. C*, vol. 120, no. 5, pp. 2485–2493, 2016.

[364] S. Huang, Q. Zhang, Y. Shiota, T. Nakagawa, K. Kuwabara, K. Yoshizawa, and C. Adachi, "Computational prediction for singlet-and triplet-transition energies of charge-transfer compounds," *J. Chem. Theory Comput.*, vol. 9, no. 9, pp. 3872–3877, 2013.

[365] T. Stein, L. Kronik, and R. Baer, "Reliable prediction of charge transfer excitations in molecular complexes using time-dependent density functional theory," *J. Am. Chem. Soc.*, vol. 131, no. 8, pp. 2818–2820, 2009.

[366] U. Salzner and A. Aydin, "Improved prediction of properties of π-conjugated oligomers with range-separated hybrid density functionals," *J. Chem. Theory Comput.*, vol. 7, no. 8, pp. 2568–2583, 2011.

[367] T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange–correlation functional using the coulomb-attenuating method (cam-b3lyp)," *Chem. Phys. Lett.*, vol. 393, no. 1-3, pp. 51–57, 2004.

[368] M. A. Rohrdanz and J. M. Herbert, "Simultaneous benchmarking of ground-and excited-state properties with long-range-corrected density functional theory," *J. Chem. Phys.*, vol. 129, no. 3, p. 034107, 2008.

[369] T. Körzdörfer, J. S. Sears, C. Sutton, and J.-L. Brédas, "Long-range corrected hybrid functionals for π-conjugated systems: Dependence of the range-separation parameter on conjugation length," *J. Chem. Phys.*, vol. 135, no. 20, p. 204107, 2011.

[370] R. Baer, E. Livshits, and U. Salzner, "Tuned range-separated hybrids in density functional theory," *Annu. Rev. Phys. Chem.*, vol. 61, pp. 85–109, 2010.

[371] T. J. Penfold, "On predicting the excited-state properties of thermally activated delayed fluorescence emitters," *J. Phys. Chem. C*, vol. 119, no. 24, pp. 13535–13544, 2015.

[372] T. Ziegler, A. Rauk, and E. J. Baerends, "On the calculation of multiplet energies by the hartree-fock-slater method," *Theor. Chim. Acta.*, vol. 43, no. 3, pp. 261–271, 1977.

[373] J. Lee, K. Shizu, H. Tanaka, H. Nomura, T. Yasuda, and C. Adachi, "Oxadiazole- and triazole-based highly-efficient thermally activated delayed fluorescence emitters for organic light-emitting diodes," *J. Mater. Chem. C*, vol. 1, no. 30, pp. 4599–4604, 2013.

[374] S. Wu, M. Aonuma, Q. Zhang, S. Huang, T. Nakagawa, K. Kuwabara, and C. Adachi, "High-efficiency deep-blue organic light-emitting diodes based on a thermally activated delayed fluorescence emitter," *J. Mater. Chem. C*, vol. 2, no. 3, pp. 421–424, 2014.

[375] I. Tamm, "Relativistic interaction of elementary particles," *J. Phys.(USSR)*, vol. 9, p. 449, 1945.

[376] S. Dancoff, "Non-adiabatic meson theory of nuclear forces," *Phys. Rev.*, vol. 78, no. 4, p. 382, 1950.

[377] S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the tamm–dancoff approximation," *Chem. Phys. Lett.*, vol. 314, no. 3-4, pp. 291–299, 1999.

[378] C. Riplinger and F. Neese, "An efficient and near linear scaling pair natural orbital based local coupled cluster method," *J. Chem. Phys.*, vol. 138, no. 3, p. 034106, 2013.

[379] J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, and H.-J. Werner, "The orbital-specific-virtual local coupled cluster singles and doubles method," *J. Chem. Phys.*, vol. 136, no. 14, p. 144105, 2012.

[380] M. E. Fornace, J. Lee, K. Miyamoto, F. R. Manby, and T. F. Miller III, "Embedded mean-field theory," *J. Chem. Theory. Comput.*, vol. 11, no. 2, pp. 568–580, 2015.