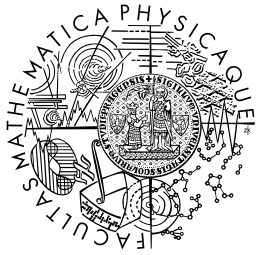

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Khanh Ngoc Pham

Development of methods for an accurate description of cohesive properties of molecular solids

Department of Chemical Physics and Optics

Supervisor: Mgr. Jiří Klimeš, Ph.D.

Study programme: Biophysics, Chemical and Macromolecular Physics

Study branch: Physics

Prague 2024

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ....

Author's signature

Acknowledgements

First and foremost, I would like to express my sincere thanks to my supervisor Mgr. Jiří Klimeš, Ph.D., for his guidance, support, motivation and patience at every stage of my PhD study and research. Thank you for warmly welcoming me into your group and for introducing me to such interesting research. It has been great being able to work with you in the ERC funded project. You have always been open to a quick chat about anything, and I have learned a lot of knowledge and experience from your lectures and our casual discussions.

Besides my supervisor, I would like to extend my special thanks to Dr. Marcin Modrzejewski for his guidance, comments and suggestions which were really influential in shaping the results of my thesis. I found it really helpful to discuss my results with you, and you have always been available and gracious in answering and explaining any questions I had about research.

I would like to thank all the members in my group and the administrators of my project. Your kind assistance and support have significantly contributed to making my study and life in Prague both easier and truly wonderful. I also want to thank to all my friends for consistently being here with me, whenever I need any help or someone to talk to.

Last but not least, I would like to express my heartfelt thanks to my mother and my sisters for supporting me all the time. Especially, my mother did everything in her power to give me opportunity to dedicate myself to study.

Title: Development of methods for an accurate description of cohesive properties of molecular solids

Author: Khanh Ngoc Pham

Department: Department of Chemical Physics and Optics

Supervisor: Mgr. Jiří Klimeš, Ph.D., Department of Chemical Physics and Optics

Abstract: Reliable prediction of the structure and stability of molecular crystals and their polymorphs is essential for understanding their properties and potential applications. However, obtaining reliable binding energies of molecular solids requires using high level electronic structure methods and strict convergence with numerical parameters. This becomes particularly challenging for molecular solids with many atoms in the unit cell, where calculations can become prohibitively expensive for high levels of theory, such as coupled clusters with singles, doubles, and perturbative triples (CCSD(T)). In this thesis, we focus on the development and assessment of approximate theoretical methods for calculation of binding energies and use four high-dispersion solids as test systems: ethane, ethylene, and orthorhombic and cubic forms of acetylene. To begin, we compare the efficiency of periodic boundary conditions (PBC) and many body expansion (MBE) approaches in calculating the binding energy of the considered systems. We discuss in detail how difficult it is to reach converged binding energy values with respect to the numerical parameters and then compare the results obtained from both approaches. In the remaining part of the thesis, we use the MBE results to examine the accuracy of random phase approximation (RPA) and Møller-Plesset (MP) perturbation theory in describing individual contributions and total binding energy of the considered systems by comparing them with reference CCSD(T) data. Additionally, we show how the accuracy of RPA can be affected by the orbital inputs used for the calculations. We find that RPA with additional corrections is promising, and its accuracy depends on the individual contributions and orbitals considered. Finally, we examine the accuracy of the proposed correction scheme, which can be used to obtain reference binding energies of the considered systems at a lower cost than the CCSD(T) method.

Keywords: Molecular solids, binding energy, periodic boundary conditions, many body expansion, random phase approximation

Název práce: Vývoj metod pro přesný popis vazebných vlastností molekulárních krystalů

Autor: Khanh Ngoc Pham

Katedra: Katedra chemické fyziky a optiky

Vedoucí bakalářské práce: Mgr. Jiří Klimeš, Ph.D., Katedra chemické fyziky a optiky

Abstrakt: Spolehlivá předpověď struktury a stability molekulárních krystalů a jejich polymorfů je nezbytná pro pochopení jejich vlastností a potenciálních aplikací. Získání spolehlivé vazebné energie molekulárních pevných látek však vyžaduje použití přesných metod elektronové struktury a striktní konvergenci s numerickými parametry. To je obzvlášť náročné u molekulárních pevných látek s mnoha atomy v jednotkové buňce, pro které mohou být výpočty nesmírně časově náročné pokud je použita vysoce přesná teoretická metoda, jako např. spřažené klastry s jedno-, dvoj- a perturbativními tříčásticovými excitacemi (CCSD(T)). V této práci se zaměřujeme na vývoj a hodnocení přibližných metod schopných spolehlivě popsat vazebnou energii přičemž pro testy jsme použili čtyři molekulární krystaly vázané hlavně disperzními silami: ethan, ethylen a ortorombickou a kubickou formu acetylenu. Nejprve porovnáváme efektivitu výpočtů vazební energie při použití periodických okrajových podmínek (PBC) a pro mnohočásticovou expanzi (MBE). Podrobně diskutujeme, jak obtížné je dosáhnout konvergovaných hodnot vazební energie s ohledem na numerické parametry, a poté porovnávme výsledky získané z obou přístupů. Ve zbytku práce používáme výsledky získané pomocí MBE ke studiu přesnosti aproximace náhodné fáze (RPA) a Møller-Plessetovy (MP) poruchové teorie. Výsledky těchto metod pro individuální příspěvky a celkové vazební energie uvažovaných systémů porovnáváme s referenčními daty získanými pomocí CCSD(T). Dále ukazujeme, jak může být přesnost RPA ovlivněna volbou orbitalů použitých pro výpočty. Zjišťujeme, že RPA s dalšími korekcemi je slibná a analyzujeme závislost přesnosti individuálních příspěvků na zvažovaných orbitalech. Nakonec zkoumáme přesnost navrhovaného korekčního schématu, které lze použít k získání referenčních vazebních energií uvažovaných systémů s nižšími výpočetními náklady než metoda CCSD(T)

Klíčová slova: Molekulární krystaly, vazebná energie, periodické okrajové podmínky, mnohočásticová expanze, přiblížení náhodné fáze

Contents

Introduction **3**

1 Theoretical and computational methods **5**
1.1 Quantum mechanics . . . 5
1.1.1 Schrödinger equation . . . 5
1.1.2 Born-Oppenheimer approximation . . . 6
1.1.3 Many electron problem and Slater determinant . . . 6
1.2 Hatree-Fock theory . . . 7
1.3 Post-Hatree-Fock methods . . . 9
1.3.1 Configuration interaction . . . 9
1.3.2 Many-body perturbation theory . . . 10
1.3.2.1 Rayleigh–Schrödinger perturbation theory . . . 10
1.3.2.2 Møller-Plesset perturbation theory . . . 12
1.3.3 Coupled cluster . . . 13
1.4 Density functional theory . . . 14
1.4.1 Electron density . . . 14
1.4.2 Hohenberg-Kohn Theorems . . . 17
1.4.3 Kohn-Sham method . . . 18
1.4.4 Exchange-correlation functionals . . . 19
1.4.5 Density functional theory with dispersion corrections . . . 21
1.5 Random phase approximation . . . 23
1.5.1 Derivation of direct RPA equation . . . 23
1.5.2 Beyond-RPA methods . . . 26
1.6 Basis sets . . . 27
1.7 Computational techniques . . . 29
1.7.1 Periodic systems . . . 29
1.7.2 Approaches for calculations of binding energy of molecular solids . . . 31
1.7.3 Correction scheme . . . 32
1.8 Methods for calculations of binding energy of molecular solids . . . 33
1.9 Summary . . . 34

2 Computational details **36**
2.1 Systems . . . 36
2.2 Orbitals used for RPA calculations . . . 37
2.3 Computational techniques . . . 37

3 Convergence of energy components with numerical parameters **39**
3.1 Many body expansion . . . 39
3.1.1 Basis set . . . 39
3.1.1.1 Two-body terms . . . 39
3.1.1.2 Non-additive terms . . . 43
3.1.1.3 Used basis sets and uncertainties . . . 46
3.1.2 Cut-off distance . . . 47
3.1.2.1 Two-body terms . . . 48

3.1.2.2 Non-additive terms . . . 51
3.1.2.3 Uncertainties . . . 52
3.1.3 Summary . . . 55
3.2 Periodic boundary conditions . . . 55
3.2.1 k-points and cell volume . . . 56
3.2.2 Plane wave basis sets . . . 60
3.2.3 PAW potentials . . . 62
3.2.4 Summary . . . 63

4 Comparison between MBE and PBC results 64
4.1 Mean field energies . . . 64
4.2 Correlation energies . . . 66
4.3 Summary . . . 67

5 Effect of orbitals on energy components in RPA calculations 68
5.1 Mean field energies . . . 68
5.2 Correlation energies . . . 70
5.3 Summary . . . 72

6 Accuracy of examined approximate methods 73
6.1 Two-body terms . . . 73
6.2 Three-body terms . . . 77
6.3 Four-body terms . . . 81
6.4 Total binding energies . . . 83
6.5 Correction scheme . . . 85
6.6 Dispersion-corrected DFT methods . . . 87

Conclusion 89

Bibliography 91

List of Figures 114

List of Tables 116

List of Abbreviations 120

List of Publications 121

A Attachments 122
A.1 Chapter 4 . . . 122
A.2 Chapter 5 . . . 126
A.3 Chapter 6 . . . 129

Introduction

Molecular solids have received considerable research interest since they constitute an important class of materials with various applications in scientific and industrial fields, especially in pharmaceuticals and material science [1, 2, 3, 4, 5]. For instance, many commonly used drugs today are typically found in solid form, such as Aspirin, Paracetamol, and Pyrazinamide. Molecular solids often exist in various polymorphs, which have same chemical composition but different molecular packing depending on crystallization conditions. In fact, at least 50% of organic molecules are estimated to exhibit polymorphism [6]. Moreover, the inclusion of cocrystals, solvates, hydrates, and salts can also introduce a wide variety of crystal structures for a single compound [7, 8, 9]. These structural variations significantly influence the crystal's characteristics, such as solubility, stability, bioavailability, and other key properties, all of which are important in the development of their practical applications [6, 10, 11, 12].

Predicting reliably structures and stabilities of molecular crystals and their polymorphs is essential for understanding their properties and potential applications [6, 13, 14]. However, there has been a shortage of valuable reference data on binding energies of molecular solids. This is because the accurate calculation of this quantity is challenging, primarily due to the need to reliably model electron correlations in periodic systems [15, 16]. Furthermore, the energy differences between polymorphs of a crystal are typically small, often on the order of one percent of the lattice energy, due to the subtle balance among different kinds of noncovalent interactions in crystal. As an example, it was observed that the difference in lattice energy between different polymorphs is less than 2 kJ/mol for ice while the binding energy is around 60 kJ/mol [16]. This makes predicting the binding energies and, in particular, the energy differences between polymorphs even more challenging [14, 17, 18, 19, 20].

Currently, binding energies of molecular solids can be obtained by two main approaches: using periodic boundary conditions (PBC) [16, 21, 22, 23] and many-body expansion (MBE) [17, 24, 25, 26, 27, 28]. The first approach appears more straightforward, but benchmark methods, such as couple cluster (CC), can not be directly used in this scheme due to their high computational requirements, except for very small systems [29]. However, coupled cluster calculations are feasible using the second approach, which decomposes the binding energy into interactions of dimers, trimers, tetramers, etc. [30, 31]. In both approaches, the binding energies depend on several parameters that need to be considered to reach high precision [21, 23]. Understanding the convergence of the binding energies with respect to these numerical parameters is important, and is one of the goals of this thesis.

In reality, one needs to use highly accurate methods, such as coupled-cluster singles, doubles, and perturbative triples [CCSD(T)], and ensure strict convergence with numerical parameters to obtain reliable binding energies of molecular solids. However, this seems impractical for all molecular solids, except for very small systems [30]. Therefore, the development of affordable methods which can describe reliably the binding energies of molecular solids is a pressing need and the main focus of this thesis. There is a range of approximate methods with differ-

ent costs and levels of accuracy currently used. Density functional theory (DFT) approximate methods are cheap but not accurate in describing long-range correlation interactions [32]. Dispersion-corrected DFT methods have been developed to address this problem [33, 34, 35, 36], but their accuracy is not always satisfactory [21, 37, 38]. The next promising methods for calculating the binding energies of molecular solids are those based on perturbation theory, such as random phase approximation (RPA) or Møller–Plesset (MP) theory [21, 23, 39, 40, 41, 42]. Moreover, recent studies indicate that the choice of orbital inputs significantly affects the accuracy of RPA calculations [43, 44, 45]. Understanding how the accuracy of RPA methods changes with different orbitals for describing the binding energies of molecular solids is important and represents another goal of this thesis.

To understand the various aspects of applying theoretical methods to molecular solids, we selected four high-dispersion molecular solids with different electrostatic moments: ethane, ethylene, and two forms of acetylene. We develop a framework based on CCSD(T) to obtain benchmark data for n-body contributions of MBE to the binding energy of the selected systems. Two main groups of methods were examined: RPA and MP theory. We start to analyze the convergence of each energy component in different methods obtained from both MBE and PBC approaches with numerical parameters (see Chapter 3). This allows us to understand the challenges in achieving converged binding energy values and identify strategies to reduce computational time. Following this, we compare the results between MBE and PBC approaches to enhance our understanding of their performance and precision (as detailed in Chapter 4). We then evaluate the effect of different orbitals on the results of each energy component in the RPA calculations obtained from MBE approach (refer to Chapter 5). This helps us to understand and identify the best orbitals for the RPA methods in describing n-body contributions to the binding energy of the considered systems. Finally, we examine the accuracy of approximate methods, including second-order Møller–Plesset perturbation theory (MP2), third-order Møller–Plesset perturbation theory (MP3), and RPA both without and with additional corrections, in describing the n-body contributions and total binding energy of molecular solids by comparing with the reference CCSD(T) data (as discussed in Chapter 6).

1. Theoretical and computational methods

In this chapter, we will give an introduction to theory and methods used to study matter at molecular level. The basic idea of most quantum mechanic methods is to solve Schrödinger equation for a system consisting of many electrons and nuclei. This equation is simple and exact in principle, but solving it exactly for systems with more than one electron is currently impossible and therefore we first discuss approaches used to solve it approximately. We then provide an overview of density functional theory which may offer computational efficiency compared to post-Hatree-Fock methods and yield very satisfactory results in many cases. Following this, we describe random phase approximation method that can offer some advantages to overcome several limitations of standard DFT functionals. We next discuss various types of basis sets used in studies of this thesis. Finally, we turn our attention to discussing approaches and theoretical methods currently used to calculate binding energy of molecular solids.

1.1 Quantum mechanics

1.1.1 Schrödinger equation

In atomic units, Hamiltonian operator for a molecular system consisting of N electrons and M nuclei can be written as:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}, \quad (1.1)$$

where M_A is ratio of the mass of nucleus A to the mass of an electron, Z_A is atomic number of nucleus A, r_{ij} is distance between i-th and j-th electrons, and r_{iA} is distance between i-th electron and A-th nucleus. The first two terms describe respectively kinetic energy of electrons and nuclei, the third term represents electrostatic interaction of electrons with nuclei, the fourth and fifth terms describe interactions between electrons, and between nuclei, respectively. Relativistic effects are not included in the Hamiltonian presented in equation 1.1 because these effects are almost entirely negligible for systems and properties studied in this thesis. The exception is spin, for which we use the standard approach of having spin up and down electrons. We note that spin-orbit interaction can lead to relativistic effects, but its magnitude is smaller than any terms in the spin-less Hamiltonian. However, these effects become important in study of heavy elements, and one has to solve Dirac's equation instead of Schrödinger equation to include them.

In quantum mechanics, the state of a system is determined by a wave function that depends on spatial coordinates of all particles and time. In this thesis, we are more interested in stationary states, i.e., those whose properties do not vary with time. Then the wave function that contains all information of a system is solution of a time-independent Schrödinger equation

$$\hat{H}\Psi(r,R) = E\Psi(r,R), \tag{1.2}$$

where r and R are the coordinates of all the electrons and nuclei, respectively. This is an eigenvalue equation with E representing the energy of the system.

1.1.2 Born-Oppenheimer approximation

The wavefunction of a system is a complicated object being a function in a $N+M$ dimensional space. This makes solving the time-independent Schrödinger equation for a system of interacting particles challenging and usually impossible. However, in many cases, electrons move faster than nuclei by several orders. This means that electrons can instantaneously respond to any changes in nuclear configuration. Therefore, one can reduce the complexity by assuming independence of the nuclear and electron wavefunctions. This is known as Born-Oppenheimer (BO) approximation. Based on this approximation, we are left with electronic Hamiltonian describing the motion of N electrons in the field of M point charges of nuclei:

$$\hat{H}_{\text{el}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}. \tag{1.3}$$

The solution of Schrödinger equation involving the electronic Hamiltonian $\hat{H}_{\text{el}}\Psi(r,R) = E_{\text{el}}(R)\Psi(r,R)$ is the electronic wave function that depends explicitly on the coordinates of electrons r, but depends parametrically on the coordinates of nuclei R. Repeatedly solving the electronic Schrödinger equation at different nuclear coordinates allows us to generate a potential energy surface (PES). This surface helps us to understand various molecular properties, such as stability, bond formation, reaction pathways, and molecular vibrations. The Born-Oppenheimer approximation is generally reliable, yet its accuracy can diminish when dealing with the systems where two electronic states have similar energy levels [46].

1.1.3 Many electron problem and Slater determinant

Molecules and solids involve many interacting particles, and we have to seek the solutions to the time-independent Schrödinger equation for N-electron wavefunction with respect to a specified position of the nuclei. However, the number of particles, combined with correlations between the particles, makes the electronic Schrödinger's equation for many-electron systems impossible to solve exactly. To find approximate solutions to the Schrödinger equation, we introduce single-electron wave functions, called orbital functions, or just simply orbitals. They take only a single particle's position as input and describe the particle's state. Let us now define a spin orbital as a wave function of a single electron describing both its spatial distribution and its spin ($\chi(x)$ with $x = (r,\alpha)$). In Hartree approach, each electron occupies its own orbital and move independently in the field created by nuclei and other electrons. Then the total wavefunction of the system with N electrons is taken as the product of the single-particle orbitals:

$$\chi(x_1,\ldots,x_N) = \chi_1(x_1)\chi_2(x_2)\ldots\chi_N(x_N). \tag{1.4}$$

The Hartree product seems to be a natural way to express the N-electron wave function. However, it has two significant limitations, that can not be ignored. The first one is that this wavefunction is uncorrelated. This deficiency stems from the definition of the Hartree product, wherein the probability of finding electron 1 in $\mathrm{d}r_1$ around r_1 remains unaffected by the positions of other electrons in the system, despite the well-known fact that the electrons repel each other to avoid the regions occupied by other electrons. The second one is that it does not fulfill antisymmetry principle, which states that the wave function must be antisymmetric when we change the position of two particles. This problem can be fixed by introducing a new form of the wave function, namely Slater determinant (SD), which is a combination of Hartree products. The Slater determinant for two particles wavefunction can be written as:

$$\Psi_{\mathrm{SD}}(x_1, x_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) \\ \chi_1(x_2) & \chi_2(x_2) \end{vmatrix} = \frac{1}{\sqrt{2}} \left[\chi_1(x_1)\chi_2(x_2) - \chi_1(x_2)\chi_2(x_1)\right]. \quad (1.5)$$

Then the wavefunction of a system of N particles can simply be written as determinant of $N \times N$ Slater matrix:

$$\Psi_{\mathrm{SD}}(x_1, ..., x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \dots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \dots & \chi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_N) & \chi_2(x_N) & \dots & \chi_N(x_N) \end{vmatrix}. \quad (1.6)$$

The Slater determinant is not a complete representation of the true wave function. However, it provides a starting point for solving the Schrödinger equation. This approach will be discussed in terms of the HF theory below.

1.2 Hatree-Fock theory

Before presenting the HF method, we mention an important result called variational principle. It states that the ground-state energy, E_0, is always less than or equal to the expectation value of $\hat{H}$ calculated with a trial wavefunction, that is:

$$\langle \Psi_{\mathrm{SD}} | \hat{H} | \Psi_{\mathrm{SD}} \rangle \geq E_0. \quad (1.7)$$

Beginning with Ψ_{SD} as a trial wave function, then varying it until the expectation value of $\hat{H}$ is minimized, we can finally obtain approximations to the wavefunction and the energy of the ground-state.

The idea of the HF theory is to replace the many-electron Schrödinger equation with a series of one-electron equations, in which each electron encounters the effects of other electrons in an average way [47]. Then the expression for the HF energy can be formulated in relation to the expectation values stemming from one-electron operator $\hat{h}$, including contributions from electron kinetic energy and interaction energy of electrons with nuclei, and from two-electron operators, including coulomb operator $\hat{J}$ and exchange operator $\hat{K}$ as

$$
\begin{aligned}
E_{\mathrm{HF}} &= \sum_{i=1}^{N} \langle \chi_i(x_1) | \hat{h}_i | \chi_i(x_1) \rangle \\
&+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\langle \chi_i(x_1) | \hat{J}_j(x_2) | \chi_i(x_1) \rangle - \langle \chi_i(x_1) | \hat{K}_j(x_2) | \chi_i(x_1) \rangle \right).
\end{aligned}
\tag{1.8}
$$

The one-electron operator $\hat{h}$, coulomb operator $\hat{J}$, and exchange operator $\hat{K}$ are defined as follows:

$$
\hat{h}_i = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}}, \tag{1.9}
$$

$$
\hat{J}_j(x_2) | \chi_i(x_1) \rangle = \langle \chi_j(x_2) | \frac{1}{r_{12}} | \chi_j(x_2) \rangle | \chi_i(x_1) \rangle, \tag{1.10}
$$

$$
\hat{K}_j(x_2) | \chi_i(x_1) \rangle = \langle \chi_j(x_2) | \frac{1}{r_{12}} | \chi_j(x_1) \rangle | \chi_i(x_2) \rangle. \tag{1.11}
$$

We then can derive the HF equations which help us to determine the optimal choice of spin orbitals. Consequently, the optimal energy value can be obtained. Details of this process are provided in Ref. [48]. Here we just show the final HF equations written as follows:

$$
\hat{F}_i \chi_i = \epsilon_i \chi_i, \quad i = 1, 2, \ldots, N. \tag{1.12}
$$

In equation 1.12, $\hat{F}_i$ is Fock operator and ϵ_i is energy of i-th spin orbital. The HF equations simplify the many-electron Schrödinger equation into a set of one-electron equations. However, each equation still depends on other $(N-1)$ spin orbitals which means that the solution should be known to solve the equations. The self-consistent field (SCF) method is employed to solve this problem. This approach starts with an initial guess for all the spin orbitals. Using this initial guess, the HF equations are solved to generate a new set of spin orbitals. These new spin orbitals are then used to recalculate the Fock operator, and the procedure is repeated until the spin orbitals reach convergence.

The HF equations are integro-differential equations, making them difficult to solve directly. In order to obtain practical results, we often use a basis set expansion to express the unknown spatial orbitals in terms of pre-defined functions. Let us introduce a set of K known functions, typically chosen as basis functions that can effectively represent molecular orbitals. Then the orbitals are approximated as linear combinations of basis functions $\varphi_\mu(r)$ with some expansion coefficients $C_{\mu i}$:

$$
\Psi_i(r) \approx \sum_{\mu=1}^{K} c_{\mu i} \varphi_\mu(r), \quad i = 1, 2, \ldots, K. \tag{1.13}
$$

Substituting this spatial orbital into equation 1.12 we obtain:

$$
\hat{F}(r) \sum_{\nu=1}^{K} c_{\nu i} \varphi_\nu(r) = \epsilon_i \sum_{\nu=1}^{K} c_{\nu i} \varphi_\nu(r), \quad i = 1, 2, \ldots, K. \tag{1.14}
$$

Multiplying by $\varphi_\mu^*(r)$ from the left leads to:

$$\sum_{\nu=1}^{K} c_{\nu i}\varphi_\mu^*(r)\hat{F}(r)\varphi_\nu(r) = \epsilon_i \sum_{\nu=1}^{K} c_{\nu i}\varphi_\mu^*(r)\varphi_\nu(r), \quad i = 1, 2, \ldots, K. \tag{1.15}$$

Then Roothaan equations are obtained by integrating both sides of equation 1.15

$$\sum_{\nu=1}^{K} F_{\mu\nu}c_{\nu i} = \epsilon_i \sum_{\nu=1}^{K} S_{\mu\nu}c_{\nu i}, \quad i = 1, 2, \ldots, K. \tag{1.16}$$

Finally, we can express closed-shell HF equations in a compact matrix form as follows:

$$\boldsymbol{FC} = \boldsymbol{SC}\boldsymbol{\epsilon}. \tag{1.17}$$

where $\boldsymbol{F}$ is Fock matrix, $\boldsymbol{C}$ is coefficient matrix, $\boldsymbol{S}$ is overlap matrix, and $\boldsymbol{\epsilon}$ is diagonal matrix of orbital energies. Due to the dependence of the Fock matrix on its own solutions, the Roothaan equations must be solved iteratively. This iterative procedure is also known as a SCF method because the SCF convergence is only reached when the new Fock operator is identical to the previous one.

1.3 Post-Hatree-Fock methods

The HF theory treats interactions between electrons in an average way, neglecting instantaneous electron-electron interactions. Thus, the energy calculated using the HF theory is higher than the true energy of the system. The difference between the true energy, E_0, and the HF energy is known as correlation energy

$$E_{\text{corr}} = E_{\text{HF}} - E_0. \tag{1.18}$$

To obtain more reliable energies, post-HF methods, such as configuration interaction (CI), many-body perturbation theory (MBPT), and coupled-cluster theory, are commonly used. Each of these methods has its strengths and limitations, and the choice of method depends on several factors, including system size, type of chemical system, computational resources, and desired accuracy. We will discuss them in detail below. In this thesis, a method derived from CC theory was used to calculate reference data, while a series of methods based on MBPT theory were used to examine the accuracy for description of binding energies of molecular solids.

1.3.1 Configuration interaction

Configuration Interaction improves upon HF single determinant wave function by representing the wave function as a linear combination of determinants. These determinants are generated from the HF orbitals by considering the ground state (Ψ_{HF}), formed from the N lowest energy orbitals obtained from solving Roothaan's equation, and the excitations of one or more electrons to give singly, doubly, triply, ..., excited determinants (denoted as Ψ_S, Ψ_D, Ψ_T, ..., etc.)

$$\Psi_{\text{CI}} = c_0\Psi_{\text{HF}} + \sum_S c_S\Psi_S + \sum_D c_D\Psi_D + \sum_T c_T\Psi_T + \ldots \tag{1.19}$$

Since CI is a variational method, the optimal coefficients can be determined by requiring that the energy is a minimum. With the optimal coefficients, the wave function, which is a linear combination of determinants, is defined. The energy of the system is then obtained by evaluating the expectation value of the Hamiltonian operator with respect to the wave function in equation 1.19.

CI is conceptually the simplest among the post-HF methods considered in this chapter. However, it can become computationally expensive, especially in the case of full-CI (FCI), where all possible determinants are added to the total wavefunction within a given basis set. This is because the number of determinants in FCI calculations grows exponentially with both the size of basis set and the number of electrons, making FCI calculations infeasible for large systems. In practice, one often truncates or approximates the FCI method to make it more computationally efficient. For example, CISD method is limited to single and double excitations. Nevertheless, a big problem of truncated CI methods is their size-inconsistency which means that the energy of two infinitely separated molecules is not exactly twice the energy of a single molecule. This is an undesirable property from a theoretical perspective, especially for calculation of binding energies.

1.3.2 Many-body perturbation theory

Many-body perturbation theory is an approach used to calculate the properties of systems by perturbing around a mean-field description and systematically including higher-order corrections beyond the mean-field description [47]. These corrections are organized into a series of terms, where each term represents an increasing level of electron correlation. We start to provide standard derivation of Rayleigh–Schrödinger perturbation theory, which is applicable to any quantum mechanical system. Following that, we delve into the discussion of Møller-Plesset (MP) perturbation theory using HF description as a starting point.

1.3.2.1 Rayleigh–Schrödinger perturbation theory

We consider a Hamiltonian operator comprising two components: a reference ($\hat{H}_0$) and a perturbation ($\hat{V}$), where $\hat{V}$ operator is much smaller compared to $\hat{H}_0$. The extent of perturbation is determined by a constant, λ, which can have any value between 0 and 1, the former representing the unperturbed case. The parameter-dependent Hamiltonian ($\hat{H}(\lambda) = \hat{H}_0 + \lambda\hat{V}$) introduces an eigenvalue problem that needs to be solved as follows:

$$\hat{H}|\Phi_i\rangle = (\hat{H}_0 + \lambda\hat{V})|\Phi_i\rangle = E_i|\Phi_i\rangle. \tag{1.20}$$

Let us assume that we know eigenfunctions and eigenvalues of $\hat{H}_0$. In other words, the Schrödinger equation for the reference Hamiltonian operator is solved:

$$\hat{H}_0|\Psi_i^0\rangle = E_i^0|\Psi_i^0\rangle. \tag{1.21}$$

We proceed to express the exact eigenvalues and eigenfunctions as a Taylor series with respect to the parameter λ:

$$E_i = E_i^0 + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots \tag{1.22}$$

$$|\Phi_i\rangle = |\Psi_i^0\rangle + \lambda|\Psi_i^{(1)}\rangle + \lambda^2|\Psi_i^{(2)}\rangle + \dots \tag{1.23}$$

The problem now is to express these quantities in terms of zeroth-order energies and matrix elements of the perturbation $\hat{V}$ between the unperturbated wave functions. We assume that $\langle\Psi_i^0|\Phi_i\rangle = 1$ and then multiply equation 1.23 by $\langle\Psi_i^0|$, it follows that all correction terms become orthogonal to the reference wave function:

$$\langle\Psi_i^0|\Psi_i^{(n)}\rangle = 0 \quad \text{for} \quad n = 1, 2, 3 \dots \tag{1.24}$$

By substituting equations 1.22 and 1.23 into equation 1.20 and equating coefficients of λ^n on both sides, we derive a set of equations:

$$\hat{H}_0|\Psi_i^0\rangle = E_i^0|\Psi_i^0\rangle \tag{1.25}$$

$$\hat{H}_0|\Psi_i^{(1)}\rangle + \hat{V}|\Psi_i^0\rangle = E_i^0|\Psi_i^{(1)}\rangle + E_i^{(1)}|\Psi_i^0\rangle \tag{1.26}$$

$$\hat{H}_0|\Psi_i^{(2)}\rangle + \hat{V}|\Psi_i^{(1)}\rangle = E_i^0|\Psi_i^{(2)}\rangle + E_i^{(1)}|\Psi_i^{(1)}\rangle + E_i^{(2)}|\Psi_i^0\rangle \tag{1.27}$$

$$\vdots$$

When each of these equations is multiplied by $\langle\Psi_i^0|$ and the orthogonality relation in equation 1.24 is applied, we obtain following expressions for the n-th order energies:

$$E_i^0 = \langle\Psi_i^0|\hat{H}_0|\Psi_i^0\rangle \tag{1.28}$$

$$E_i^{(1)} = \langle\Psi_i^0|\hat{V}|\Psi_i^0\rangle \tag{1.29}$$

$$E_i^{(2)} = \langle\Psi_i^0|\hat{V}|\Psi_i^{(1)}\rangle \tag{1.30}$$

$$\vdots$$

The task now is to solve the set of equations for $|\Psi_i^{(n)}\rangle$ and consequently determine the (n+1)-th order correction of energy.

Let us consider equation 1.26, which determines the first-order wave function $\langle\Psi_i^{(1)}|$. This equation can be rearranged as follows:

$$(E_i^0 - \hat{H}_0)|\Psi_i^{(1)}\rangle = (\hat{V} - E_i^{(1)})|\Psi_i^0\rangle. \tag{1.31}$$

By using the orthonormality relation $\langle\Psi_n^0|\Psi_i^0\rangle = \delta_{ni}$ and multiplying the equation 1.31 by $\langle\Psi_n^0|$ from the left, we can substitute it into the expression for the second-order energy correction (equation 1.30) and obtain:

$$E_i^{(2)} = \langle\Psi_i^0|\hat{V}|\Psi_i^{(1)}\rangle = \sum_n \langle\Psi_i^0|\hat{V}|\Psi_n^0\rangle\langle\Psi_n^0|\Psi_i^{(1)}\rangle = \sum_{n\neq i} \frac{|\langle\Psi_i^0|\hat{V}|\Psi_n^0\rangle|^2}{E_i^0 - E_n^0}. \tag{1.32}$$

One can proceed similarly to obtain the higher order energies.

1.3.2.2 Møller-Plesset perturbation theory

Møller-Plesset (MP) perturbation theory is a specific form of many-body perturbation theory, which begins with the HF systems as its unperturbed system. The unperturbed Hamiltonian is taken as the sum of Fock operators:

$$\hat{H}_0 = \sum_i f(i) = \sum_i [h(i) + v^{\mathrm{HF}}(i)], \tag{1.33}$$

where v^{HF} is the HF potential including coulomb and exchange potentials. The sum of the Fock operators counts the average the electron-electron repulsion twice. Then the perturbation which is the difference between the exact Hamiltonian $\hat{H}$ and the unperturbed Hamiltonian $\hat{H}_0$ can be expressed as:

$$\hat{V} = \sum_{i<j} r_{ij}^{-1} - \sum_i v^{\mathrm{HF}}(i). \tag{1.34}$$

The zero-order wave function corresponds to the HF determinant, and the zeroth-order energy can be described as the sum of the orbital energies

$$E_0^{(0)} = \langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle = \sum_i^N \epsilon_i. \tag{1.35}$$

The HF energy is the sum of zeroth-order and first-order energies. Thus, it is necessary to go to higher-order corrections to improve on the HF energy. Examining the general formula for the second-order energy derived in equation 1.32, it becomes clear that we need to identify the states $|\Psi_n^0\rangle$ first. According to the Brillouin theorem, singly excited Slater determinants do not contribute. Additionally, triply excited states do not interact with $|\Psi_0\rangle$ due to the two-particle character of the perturbation. As a result, our focus narrows to doubly excited Slater determinants of the form $|\Psi_{ab}^{rs}\rangle$, where two electrons are excited from occupied orbitals a and b to virtual orbitals r and s. For this particular Slater determinant, we have an eigenvalue equation in the form:

$$\hat{H}_0 |\Psi_{ab}^{rs}\rangle = (E_0^{(0)} - (\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s))|\Psi_{ab}^{rs}\rangle. \tag{1.36}$$

Finally, the second-order correction is obtained by summing all possible double excitations, subject to the conditions that we sum over all pairs of occupied orbitals (a and b) where b is greater than a, as well as all pairs of virtual orbitals (r and s) where s is greater than r

$$E_0^{(2)} = \sum_{a<b,r<s} \frac{|\langle \Psi_0 | \sum_{i<j} r_{ij}^{-1} | \Psi_{ab}^{rs} \rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s} \tag{1.37}$$

$$= \sum_{a<b,r<s} \frac{|\langle ab|rs\rangle - \langle ab|sr\rangle|^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}. \tag{1.38}$$

Higher-order corrections in Møller-Plesset perturbation theory are determined in a similar manner. Specifically, the third-order correction, as seen in the MP3 method, is calculated using doubly excited determinants. The MP methods, especially MP2, have been widely used for calculating the properties of quantum systems thanks to the balance between accuracy and computational cost. Furthermore, unlike to CI, MP methods are size-extensive, making them particularly advantageous when studying large systems. However, these methods are not variational, meaning they may not always yield an upper bound to the true ground-state energy.

1.3.3 Coupled cluster

The fundamental concept of coupled cluster methods is to include all possible corrections of wavefunction of a specific type to an infinite order [47, 49]. The electronic wave function is expanded as an exponential operator acting on a reference wave function:

$$|\Psi_{\mathrm{CC}}\rangle = e^{\hat{T}}|\Psi_0\rangle, \tag{1.39}$$

with $e^{\hat{T}} = \hat{I} + \hat{T} + \frac{1}{2}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \dots$. The cluster operator $\hat{T}$ is composed of excitation operators $\hat{T}_i$ acting on the reference determinant $|\Psi_0\rangle$ to generate all i-th excited Slater determinants

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_N, \tag{1.40}$$

$$\hat{T}_1|\Psi_0\rangle = \sum_{a}^{N_{occ}} \sum_{r}^{N_{virt}} t_a^r |\Psi_a^r\rangle, \tag{1.41}$$

$$\hat{T}_2|\Psi_0\rangle = \sum_{a<b}^{N_{occ}} \sum_{r<s}^{N_{virt}} t_{ab}^{rs} |\Psi_{ab}^{rs}\rangle, \tag{1.42}$$

$$\vdots$$

The expansion coefficients $t_{ab\dots}^{rs\dots}$ are called amplitudes. The coupled cluster method is defined by the truncation of the excitation operator $\hat{T}$. We consider first the coupled cluster with singles and doubles (CCSD) model ($\hat{T} = \hat{T}_1 + \hat{T}_2$). The Schrödinger equation can be written using the coupled-cluster wave function defined in equation 1.39 as:

$$\hat{H}e^{\hat{T}}|\Psi_0\rangle = Ee^{\hat{T}}|\Psi_0\rangle, \tag{1.43}$$

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = E|\Psi_0\rangle. \tag{1.44}$$

By multiplying the equation 1.44 with the reference and excited state determinants, we can obtain equations for the amplitudes:

$$\langle\Psi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = E, \tag{1.45}$$

$$\langle\Psi_a^r|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = 0, \tag{1.46}$$

$$\langle\Psi_{ab}^{rs}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = 0. \tag{1.47}$$

The energy can be determined by taking the Schrödinger equation as presented in equation 1.43 and projecting it onto the reference wave function:

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E\langle\Psi_0|e^{\hat{T}}\Psi_0\rangle, \tag{1.48}$$

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E\langle\Psi_0|(\hat{I} + \hat{T}_1 + \dots)\Psi_0\rangle, \tag{1.49}$$

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E. \tag{1.50}$$

This expression can then be simplified using the fact that the electronic Hamiltonian contains only one- and two-electron operators and the Brillouin theorem:

$$E = \langle\Psi_0|\hat{H}(\hat{I} + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2)|\Psi_0\rangle, \tag{1.51}$$

$$E = E_{\mathrm{HF}} + \sum_{a<b}^{N_{occ}} \sum_{r<s}^{N_{virt}} (t_{ab}^{rs} + t_a^r t_b^s - t_a^s t_b^r) \langle \Psi_0 | \hat{H} | \Psi_{ab}^{rs} \rangle, \tag{1.52}$$

$$E = E_{\mathrm{HF}} + \sum_{a<b}^{N_{occ}} \sum_{r<s}^{N_{virt}} (t_{ab}^{rs} + t_a^r t_b^s - t_a^s t_b^r)(\langle ab|rs\rangle - \langle ab|sr\rangle). \tag{1.53}$$

The procedure for deriving the CC method with higher excitations can be obtained in a similar manner. The inclusion of the $\hat{T}^3$ in the operator $\hat{T}$ leads to couple-cluster with singles, doubles and triples (CCSDT) method which becomes too computationally expensive for large systems. As an alternative approach, the triples contribution can be calculated separately using perturbation theory and then added to the CCSD results. More specifically, the triples contribution is calculated from the formula established in the MP4 method by using the CCSD amplitudes instead of the perturbation coefficients and adding a fifth-order perturbation term, which describes the coupling between singles and triples. This combined method is known as CCSD(T), which is currently considered as a "golden standard" of quantum chemistry [50, 51].

CC methods are typically not variational, but they are size-consistent, similar to MP methods. In the CC theory, it is important to note that a specific excitation level is not solely a result of a single excitation operator. Instead, there are both "connected" and "disconnected" terms that contribute to this level. For example, a coupled-cluster wave function limited to connected double excitations not only includes doubles, but also higher ones (quartets, etc.). This is in contrast with both MP and CI methods and therefore CC theory provides a better description of electron correlation effects at a given truncation level.

1.4 Density functional theory

Calculating the full wave function of a quantum system is a computationally demanding task and can pose significant challenges related to memory storage. In this section, we discuss the evaluation of the energy of a system using electron density instead of wave function. We begin by describing how the energy can be obtained from electron density. Following this, we mention two theorems postulated by Hohenberg and Kohn. Next, we turn our attention to Kohn-Sham method, which provides a practical approach to determine the energy and properties of a system via electron density. In principle, the Kohn-Sham method is conceptually simple and exact. However, it introduces a new quantity, exchange-correlation energy, and we discuss this in the following part. Finally, we discuss several approaches used to improve the accuracy of standard DFT methods.

1.4.1 Electron density

The wave function of a system with many electrons is expressed as a function of four coordinates for each electron (three spatial coordinates and one spin coordinate). Consequently, as the number of electrons (N) in the system increases, the wave function rapidly becomes very complicated, depending on $4N$ variables. Nevertheless, the energy of the system does not rely explicitly on all $4N$ variables, since it is an expectation value that can be represented by one- and two-electron

integrals. The replacement of the wave function with the electron density reduces the complexity of dealing with $4N$ variables to a simpler four-dimensional density distribution.

The electron density, denoted as $\rho(r)$, is a measure of the probability of finding an electron in the volume element $\mathrm{d}r$ at a particular position r in space [52]

$$\rho(r) = \int |\Psi(r)|^2 \mathrm{d}r. \tag{1.54}$$

The integral of the electron density over all space is equal to the total number of electrons in the system. This property ensures that the electron density is capable of describing the total number of electrons in the system. In addition, the electron density exhibits a discontinuity as the electron approaches the nucleus. This is a consequence of Coulombic attraction between the negatively charged electrons and the positively charged nucleus. When the electron approaches very close to the nucleus at R_A, the Coulomb energy diverges for $r_i = R_A$, which results in a cusp at the nucleus. This property shows that the electron density is capable of describing the position of the nucleus and the nuclear charge Z_A of the system. Therefore, the electron density truly contains all necessary information about the system, i.e., $\rho(r) \Rightarrow \{N, R_A, Z_A\}$.

The electronic Hamiltonian can then be written so that it depends only on these variables, indicating that the energy of the system can be obtained from the electron density. Let us start to consider the interaction of the electrons with nuclei which is defined as:

$$E_{\text{nucl}} = -\langle \Psi \mid \sum_{i}^{N} \sum_{A}^{M} \frac{Z_A}{|r_i - R_A|} \mid \Psi \rangle \tag{1.55}$$

$$= -\int \mathrm{d}r_1 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N) \sum_{i}^{N} \sum_{A}^{M} \frac{Z_A}{|r_i - R_A|} \Psi(r_1, \ldots, r_N). \tag{1.56}$$

The sum over the electrons ($i = 1 - N$) can be written explicitly as follows:

$$\begin{aligned} &-\int \mathrm{d}r_1 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N) \sum_{A}^{M} \frac{Z_A}{|r_1 - R_A|} \Psi(r_1, \ldots, r_N) - \cdots - \\ &-\int \mathrm{d}r_1 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N) \sum_{A}^{M} \frac{Z_A}{|r_N - R_A|} \Psi(r_1, \ldots, r_N). \end{aligned} \tag{1.57}$$

The terms in above equation are equivalent due to the presence of the wave function appearing twice in the integration and the antisymmetry of the wave function. As a result, we can combine and sum them as:

$$-N \int \mathrm{d}r_1 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N) \sum_{A}^{M} \frac{Z_A}{|r_i - R_A|} \Psi(r_1, \ldots, r_N). \tag{1.58}$$

The integration can now be divided into two parts: one involving an integral over r_1 and the other encompassing the remaining coordinates:

$$-\int \mathrm{d}r_1 \sum_{A}^{M} \frac{Z_A}{|r_1 - R_A|} N \int \mathrm{d}r_2 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N) \Psi(r_1, \ldots, r_N). \tag{1.59}$$

We can see that the term in the integration over $r_2, \ldots, r_N$ coordinates corresponds to the definition of the electron density:

$$\rho(r_1) = N \int \mathrm{d}r_2 \cdots \mathrm{d}r_N \Psi^*(r_1, \ldots, r_N)\Psi(r_1, \ldots, r_N). \tag{1.60}$$

Finally, the energy of interaction of electrons with nuclei can be obtained as follows:

$$E_{\text{nucl}} = -\int \mathrm{d}r \sum_A^M \frac{Z_A}{|r_i - R_A|}\rho(r). \tag{1.61}$$

Summing over the nuclei results in Coulomb potential $V_{\text{nucl}}(r) = -\sum_A \frac{Z_A}{|r-R_A|}$, which simplifies the energy expression to:

$$E_{\text{nucl}} = \int \mathrm{d}r V_{\text{nucl}}(r)\rho(r). \tag{1.62}$$

The above equation is just one example of a single particle local potential. However, the nuclei are considered "external" potential. This is because due to the BO approximation, the positions of nuclei are fixed and do not vary during the calculation of electronic wavefunctions and energies. Then the general expression for a local external potential $V_{\text{ext}}(r)$ can be formulated as:

$$E_{\text{ext}} = \int \mathrm{d}r V_{\text{ext}}(r)\rho(r). \tag{1.63}$$

In a similar manner, the classical Coulomb energy of the electron density can be represented as:

$$E_{\text{class}} = \frac{1}{2}\int \mathrm{d}r_1 \mathrm{d}r_2 \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|}. \tag{1.64}$$

The kinetic energy can not be easily written as a functional of the density. The earliest kinetic energy functionals, proposed by Fermi, Thomas, and others, were based on non-interacting uniform electron gas (UEG) model [52]. In this approximation, it is assumed that the electrons form a uniform (or homogeneous) electron gas, meaning that the electrons move in a background uniform distribution of positive charge. However, when applied to atoms and molecules, these functionals failed to provide accurate results. Addressing this issue has driven the development of more sophisticated and accurate kinetic energy functionals in DFT.

Another challenging component in electronic Hamiltonian is electron-electron interaction, which can not be integrated out to yield an explicit functional of the electron density. The energy associated with this term can be expressed as:

$$E_{\text{el}} = \int \mathrm{d}r_1 \mathrm{d}r_2 \frac{\gamma_2(r_1, r_2; r_1, r_2)}{|r_1 - r_2|}, \tag{1.65}$$

where $\gamma_2(r_1, r_2; r_1, r_2)$ is second order spin-less reduced density matrix (2-RDM). This is derived by integrating out all the electron (and spin) coordinates of the product $\Psi^*(r'_1, \ldots, r'_N)\Psi(r_1, \ldots, r_N)$ except for the first two.

1.4.2 Hohenberg-Kohn Theorems

The use of electron density as the main variable in electronic structure calculations has gained significant attention and development. The key motivation for this development is due to Hohenberg-Kohn theorems, established by Pierre Hohenberg and Walter Kohn [53]. These theorems are foundational principles in the field of DFT.

The original work of Hohenberg and Kohn focuses on uniform electron gas as a model system. In this context, the electronic Hamiltonian, expressed as $\hat{H}_{\text{el}} = \hat{T} + \hat{V} + \hat{U}$, has the key difference from the electronic Hamiltonian approximated by the BO approximation being that the non-universal term $\hat{V}$ accounts for the interaction of electrons with an external potential $\hat{V}_{\text{ext}}$. $\hat{T}$ and $\hat{U}$ are the kinetic operator and electron-electron interaction operator, respectively. The first Hohenberg-Kohn theorem demonstrates that, for a given N-electron system, there exists a one-to-one correspondence between the local external potential V_{ext} and the electron density $\rho(r)$ [53, 54]. Given the external potential (V_{ext}) and the number of electrons, we can solve Schrödinger's equation to obtain the wavefunction (Ψ). Then the electron density $\rho(r)$ can be obtained from the wave function by integrating over all coordinates. Therefore, the external potential uniquely determines the electron density. The total energy then can be obtained from the electron density as follows:

$$E[\rho] = \int \mathrm{d}r V_{\text{ext}}(r)\rho(r) + F[\rho], \tag{1.66}$$

in which $F[\rho] = \langle\Psi|\hat{T} + \hat{U}|\Psi\rangle$ is a general functional of the density (since $\Psi \to \rho$) that can be used with any external potential and gives us the exact kinetic and electron-electron energies.

We now want to obtain the ground state energy defined by equation 1.66. This can be based on second theorem of Hohenberg and Kohn, which states that the ground state energy can be derived from the electron density by the use of variational principle [52, 53]. The electron density, which provides a minimum of the ground state energy, is therefore the exact ground state density.

Since the wave function is a unique functional of the electron density, every trial wave function $\Psi^{'}$ corresponds to a trial density $\rho^{'}(r)$. Then the ground state energy is obtained as:

$$E_0 = \min_{\Psi^{'}} \langle\Psi^{'}|\hat{H}|\Psi^{'}\rangle \tag{1.67}$$

The original proof of this second theorem was initially established using variational calculus [53]. Then, a different approach, known as constrained-search, was introduced by Levy and Lieb [55, 56]. In this approach, the minimization can be carried out in two steps. In the first step, a trial electron density $\rho^{'}(r)$ is fixed, and the corresponding wavefunction $\Psi^{'}(r)$ that minimizes the energy for this fixed density is determined. In the second step, from the set of the densities and corresponding wavefunctions obtained in the first step, the specific combination that yields the lowest energy is selected. These steps can be expressed as follows:

$$E_0 = \min_{\rho'} \left(\min_{\Psi' \to \rho'} \langle \Psi' | \hat{T} + \hat{U} | \Psi' \rangle + \int V'_{\text{ext}}(r) \rho'(r) \, \mathrm{d}r \right) \tag{1.68}$$

$$= \min_{\rho'} \left(F[\rho'] + \int V'_{\text{ext}}(r) \rho'(r) \, \mathrm{d}r \right), \tag{1.69}$$

in which $F[\rho'] = \min_{\Psi' \to \rho'} \langle \Psi' | \hat{T} + \hat{U} | \Psi' \rangle$ is the universal functional which requires no explicit knowledge of $\rho'(r)$.

In summary, Hohenberg-Kohn theorems offer a potential approach for determining the ground state of a quantum system by considering only the electron density. However, the equation of $F[\rho]$ inherently involves the all-electron wavefunction. Consequently, obtaining an explicit formulation for $F[\rho]$ in terms of the electron density is a highly complex task and does not appear feasible for realistic systems.

1.4.3 Kohn-Sham method

As mentioned before, the kinetic energy of electrons is challenging to calculate directly in DFT because it depends on the electronic wave functions, which are not explicitly known. Kohn-Sham (KS) method addresses this challenge by introducing an auxiliary system of non-interacting electrons with the same electron density as the real system [57]. The kinetic energy can then be obtained for these auxiliary electrons and is more accurate than when obtained from the density. Overall, the KS equations contain a single particle Hamiltonian common for all the orbitals so that, in practice, one needs to solve similar equations to the HF equations. Then the total energy of a real system in KS model is expressed as:

$$E_{\text{KS}}[\rho] = T_{\text{S}}[\Psi] + E_{\text{ext}}[\rho] + J[\rho] + E_{\text{xc}}[\rho]. \tag{1.70}$$

In the equation, T_{S} is the kinetic energy of non-interacting electrons, J is the Coulomb interaction of electron density and can be calculated using electron-electron repulsion integrals. The key component of the equation 1.70 is exchange-correlation (XC) energy (E_{xc}). This term accounts for the difference between real and non-interacting kinetic energies, non-classical part of the electron-electron energy and also include self-interaction correction. The last energy arises in DFT because, in contrast to the HF approach, there exists an unphysical interaction of an electron with itself arising from the Coulomb term of the energy (in HF scheme the Coulomb and exchange terms cancel for an electron interacting with itself).

The task now is to minimize the total energy functional $E[\rho]$ under the constraint that the electron density $\rho(r)$ integrates to the number of electrons (N). This constrained minimization leads to KS equations for optimal orbitals

$$\hat{f}_{\text{KS}}(r) \Psi_i(r) = \epsilon_i \Psi_i(r), \tag{1.71}$$

in which KS operator can be expressed as follows:

$$\hat{f}_{\text{KS}}(r) = -\frac{1}{2} \nabla_r^2 + V_{\text{ext}}(r) + \hat{J}(r) + V_{\text{xc}}(r). \tag{1.72}$$

The last term in the above equation is an unknown exchange correlation potential, and can be expressed as functional derivative of XC energy with respect to the density: $\frac{\delta E_{\mathrm{XC}}[\rho(r)]}{\delta\rho(r)}$. The difference between the HF and KS equations is that the exchange integral in the Fock operator is replaced by V_{xc} in the KS operator. Similar to the HF method, the KS equations are typically solved using the SCF method and with orbitals expanded using basis functions. In principle, the KS equation is exact, meaning that for an exact expression of XC functional, the system can be solved exactly. However, an exact XC functional has not been found so far, approximate functionals must be used, which will be discussed in detail in the following section.

1.4.4 Exchange-correlation functionals

The search for better and better XC functionals is never-ending mission in DFT. Although there are some physical constraints that XC functionals should obey, there is no systematic strategy for progressively approaching exact XC functional, as in the case of wavefunction-based approaches, where the variational method is cornerstone. In an attempt to formulate an expression for a complicated XC functional, it is common practice to divide it into two components: exchange functional and correlation functional. There is currently a large number of forms of the XC functional available. These functionals are traditionally categorized into several types based on their dependence on the electron density.

The simplest approximation for the XC energy is local density approximation (LDA) assuming that the XC energy density depends only on the electron density at each point in space. The expression for the XC energy based on LDA approximation is written as:

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[\rho(r)] = \int \rho(r)\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho(r))\mathrm{d}r = \int \rho(r)\left[\left(\epsilon_{\mathrm{x}}^{\mathrm{LDA}}(\rho(r)) + \epsilon_{\mathrm{c}}^{\mathrm{LDA}}(\rho(r))\right)\right]\mathrm{d}r. \tag{1.73}$$

In LDA, the exchange part relies on the HF exchange energy of UEG, which can be evaluated analytically. Consequently, LDA gives the DFT exchange energy exactly for UEG. However, the correlation part is more complicated, and can only be obtained approximately by solving the UEG system to a high accuracy using approaches such as quantum Monte Carlo [52]. Commonly used LDA functionals are VWN5 [58], PZ81 [59], PW92 [60], and so on. LDA methods are rather accurate for certain molecular properties, such as geometric structure and vibration frequency [61, 62]. However, the binding energy of molecules and solids calculated using LDA is often significantly overestimated [63, 64, 65]. In a recent study, a new functional called meta-LDA, derived from LDA, has been formulated and demonstrates improved accuracy compared to LDA [66].

The subsequent significant breakthrough in formulating $\mathrm{E_{xc}}$ was the development of generalized gradient approximation (GGA). Similar to LDA, the GGA methods are local and depend on the electron density at a specific point. However, these methods also take into account gradient of the electron density at that point. This inclusion aims to capture the non-uniform nature of the electron density. Then the GGA XC energy can be expressed as follows:

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[\rho(r)] = \int \rho(r)\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}(\rho(r), \nabla\rho(r))\mathrm{d}r. \tag{1.74}$$

The GGA methods represent an improvement in the accuracy of DFT calculations compared to LDA methods, making them a more reliable choice for a wide range of systems and properties [65, 67]. There are many GGA functionals that have gained widespread use in DFT calculations, including PBE [68], BLYP [69], BP86 [68, 70], and many others. Among them, PBE and its modifications are the most popular functionals, especially for condensed matter physics, due to their reasonable accuracy over a wide range of systems [65, 71]. However, PBE has been also known to significantly underestimate the binding energies in many cases [72, 73, 74].

Additional improvements and modifications to GGA functionals, known as meta-GGA functionals, can be achieved by introducing a kinetic energy dependence into the GGA XC functional. Then the meta-GGA XC energy can be written as:

$$E_{\mathrm{XC}}^{\mathrm{MGGA}}[\rho(r)] = \int \rho(r)\epsilon_{\mathrm{xc}}^{\mathrm{MGGA}}(\rho(r), \nabla\rho(r), \tau(r))\mathrm{d}r, \tag{1.75}$$

where the kinetic-energy density is given by:

$$\tau(r) = \sum_i \frac{1}{2}|\nabla\Psi_i(r)|^2. \tag{1.76}$$

The inclusion of the total kinetic energy in the XC energy makes meta-GGA functionals non-local. This increased flexibility allows them to more accurately describe the properties of molecules and solids compared to LDAs and GGAs [65, 71]. However, meta-GGA functionals typically exhibit greater sensitivity to integration grid compared to GGAs, especially when dealing with weakly interacting systems [75, 76]. The meta-GGA functionals can be categorized into two main groups: empirical and non-empirical. The functionals like MXX series (M05, M06, M08, M11, etc.) [77, 78, 79, 80, 81, 82, 83, 84] were developed through empirical fitting of parameters. The popular non-empirical functionals include PKZB [85], TPSS [86], revTPSS [87], MSX (X=0,1,2) [88], SCAN [89], TM [90], and HLE16 [91]. Among them, SCAN has emerged as a promising choice for accurately describing properties of various systems [92, 93, 94, 95]. However, the application of SCAN functional is limited by its sensitivity to the density of the numerical integration grid [96]. To improve this numerical stability, rSCAN functional was developed while retaining the accuracy of original SCAN funtional [96]. Nevertheless, extensive tests show that rSCAN yields large errors for some properties, such as heats of formation and atomization energies [97, 98]. This is because rSCAN violates some of the constraints that the original SCAN satisfies. The r^{2}SCAN functional was then developed to address these issues [99], and it performs well for many properties of molecules and solids [100, 101, 102, 103]. In recent years, density-corrected SCAN functional has been also used to alleviate the shortcoming of SCAN functional [104, 105, 106].

The inclusion of each additional ingredient into XC functionals described above offers a systematic improvement in the accuracy. However, these XC functionals still encounter three main issues: self-interaction error [59, 107], long-range correlation interactions [32], and strong correlation effects [108]. A simple way to reduce the self-interaction error is to use an exchange functional, which combines contributions from the HF exchange $E_{\mathrm{X}}^{\mathrm{HF}}$ with the $E_{\mathrm{X}}^{\mathrm{LDA}}$ or $E_{\mathrm{X}}^{\mathrm{GGA}}$ component. This way results in functionals known as hybrids, which can be expressed as

follows:

$$E_{\mathrm{xc}}^{\mathrm{Hybrids}} = c_{\mathrm{x}} E_{\mathrm{x}}^{\mathrm{HF}} + (1 - c_{\mathrm{x}}) E_{\mathrm{x}}^{\mathrm{DFT}} + E_{\mathrm{c}}^{\mathrm{DFT}}. \tag{1.77}$$

It is important to note that the inclusion of full exact exchange may yield results that are even less accurate than those obtained with GGA methods. This is because full exact exchange is not compatible with the correlation terms used in GGA functionals. Therefore, the choice of the fraction of exact exchange (c_{x}) in hybrid functionals is a crucial consideration. Some hybrid functionals widely used include B3LYP [69, 109], PBE0 [110], and SCAN0 [111]. Clearly, hybrid functionals are useful for the systems with strong-self-interaction errors and also for atomization energies. However, they can encounter limitations when applied to extended systems, because in reality the exchange interactions in extended systems are screened at long distances, which are not adequately accounted for by standard correlation functionals in hybrid functionals. To address this problem, range-separated hybrid functionals, which split the exchange energy into short-range and long-range components and treat them differently, are usually used [112, 113, 114, 115, 116]. In this approach, HF is usually used to evaluate the short-range interactions, while semilocal DFT is applied to treat the long-range interactions. Another advancement in hybrid functionals is double hybrid (DH) theory, which combines standard hybrid functionals with second-order wave function methods. Typically, a hybrid calculation is performed first, and then an MP2 correction, evaluated on the DFT orbitals, is added to the XC energy. Many DH functionals have been proposed [111, 117, 118, 119], and their performance is excellent for many properties of different materials [120, 121, 122].

1.4.5 Density functional theory with dispersion corrections

As mentioned above, the dispersion interactions can not be accurately described by local and semi-local DFT functionals. In other words, standard DFT functionals can not provide the correct $-C_6/R^6$ dependence, which is crucial for describing the dispersion interaction energy at large interatomic distances. Despite extensive efforts to develop local functionals for modeling dispersion interactions, these attempts have proven to be unsuccessful [123, 124]. Therefore, there has been an interest in methods for capturing dispersion interactions within DFT over the last three decades. There are several ways to do this. Here, we only focus on discussing the method we used in our work, usually called DFT-D or sometimes DFT+disp. In this method, the calculations of missing dispersion interactions are based on an atom pairwise additive treatment. The general form for the dispersion energy, which is simply added to the total DFT energy, is:

$$E_{\mathrm{disp}}^{\mathrm{DFT+disp}} = -\sum_{\mathrm{AB}} \sum_{n=6,8,10,\ldots} s_n \frac{C_n^{\mathrm{AB}}}{R_{AB}^n} f_{\mathrm{damp}}(R_{\mathrm{AB}}). \tag{1.78}$$

In the formula, the sum is over all atom pairs in the system, R_{AB} represents the distance between atoms A and B, and C_n^{AB} corresponds to the averaged n-th order dispersion coefficient. Some studies have demonstrated that using only the C_6 coefficient is insufficient for describing medium and short-range dispersion interactions [125, 126, 127]. Therefore, it is necessary to incorporate higher

coefficients, such as C_8 and C_{10}, as they contribute significantly in equilibrium regions [127, 128]. However, the considerably large values of these coefficients can lead to the amplification of their corresponding errors, thereby introducing instability in the correction [125]. The global scaling factors s_n are commonly employed to adjust the correction to the repulsive behavior of the chosen density functional [129]. The function $f_{\text{damp}}(R_{AB})$ is a damping function which is used to avoid singularities for small R and double-counting effects of correlation at intermediate distances. Finally, non-additive dispersion effects arising from three-body interactions can be calculated by adding C_9, or Axilrod–Teller–Muto (ATM) contribution [130, 131].

Among the various DFT+disp methods, the ones developed by Stefan Grimme are widely popular. The relatively early DFT-D1 version [129] was designed for modeling noncovalent interactions for biomolecules [132]. DFT-D2, an advancement upon DFT-D1, modifies the approach by including damping functions to the C_6/R^6 terms to provide more accurate description of dispersion interactions [133]. DFT-D3 introduces a slightly higher level of complexity than DFT-D2 by including additional ingredients, such as the coefficients for eighth-order dispersion terms and a new set of cutoff radii [125]. Additionally, the parameters C_6 and C_8 in DFT-D3 version depend on the coordination number of each atom, which are fixed in older versions. This makes DFT-D3 more widely used compared to the older versions for various materials [33, 134, 135]. A similar method, named as DFT-D3BJ, was developed later by using Becke-Johnson damping function to improve the accuracy of dispersion interactions, particularly at short distances [136, 137]. The newest variant, DFT-D4 [138], continues to add accuracy and complexity by taking into account atomic partial charges and their influence on the dispersion parameters. By doing so, DFT-D4 captures the dependency of each atom's dispersion parameters on its surrounding environment, leading to a more accurate and adaptable representation of dispersion forces [139, 140, 141]. There are other dispersion correction methods which have also gained popularity, including dipole-exchange hole model (XDM) [142, 143, 144], Tkatchenko-Scheffler (TS) model [145], and many-body dispersion (MBD) [146] and its rsMBD [147], and (uMBD) [148] variants, nonlocal many-body dispersion (MBD-NL) [149]. The accuracy of these models has been examined for many properties of different materials [36, 150, 151, 152, 153].

Although the DFT+disp methods have been widely employed to enhance the accuracy of standard DFT methods, they still have several limitations. One such limitation is their difficulty in accurately describing systems with significant changes in electron density, as the interaction in these methods does not directly derive from the electron density. Moreover, the dispersion contribution (E_{disp}) in DFT+disp methods is inherently attractive, which might result in a tendency to overbind binding energies of molecules and solids. To address this issue, the DFT+disp methods are recommended to be used with the DFT functionals that do not overbind. In other words, the choice of the DFT functional is critical to the overall accuracy of the DFT+disp methods.

1.5 Random phase approximation

Random phase approximation is an approximate electronic structure method, which combines the elements of DFT and wave-function theory (WFT). From the DFT perspective, it is an approach at fifth rung of Jacob's ladder for the exchange-correlation approximations [154], which uses not only electron density, but also virtual orbitals and orbital energies of Kohn-Sham system [155]. From the wave-function point of view, RPA is a ring-only approximation of coupled-cluster doubles wave function method with HF or KS reference mean-field Hamiltonian [156].

1.5.1 Derivation of direct RPA equation

RPA can be derived from several theoretical frameworks, including adiabatic-connection fluctuation-dissipation (ACFD) theorem [157, 158, 159, 160], many-electron Green's function theory [160, 161], and coupled cluster theory [49, 156]. Here we only present the process of deriving RPA correlation energy using ACFD theorem.

Let us first use adiabatic connection (AC) technique to obtain the ground-state total energy of an interacting many-body Hamiltonian [162]

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda \hat{H}_1(\lambda), \tag{1.79}$$

where λ is a coupling parameter that connects a reference Hamiltonian $\hat{H}_0 = \hat{H}(\lambda = 0)$ with many-body Hamiltonian $\hat{H} = \hat{H}(\lambda = 1)$, $\hat{H}_1(\lambda)$ is perturbative Hamiltonian.

The $\hat{H}_0$ and $\hat{H}_\lambda$ Hamiltonians for the electronic systems have the following forms:

$$\hat{H}_0 = \sum_{i=1}^{N} \Big[-\frac{1}{2}\nabla_i^2 + V_{\lambda=0}^{\text{ext}}(i) \Big]. \tag{1.80}$$

$$\hat{H}(\lambda) = \sum_{i=1}^{N} \Big[-\frac{1}{2}\nabla_i^2 + V_{\lambda}^{\text{ext}}(i) \Big] + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{\lambda}{|r_i - r_j|}. \tag{1.81}$$

By substituting equations 1.80 and 1.81 into equation 1.79, we obtain the perturbative Hamiltonian $\hat{H}_1(\lambda)$:

$$\hat{H}_1(\lambda) = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|r_i - r_j|} + \frac{1}{\lambda}\sum_{i=1}^{N} \Big[V_{\lambda}^{\text{ext}}(i) - V_{\lambda=0}^{\text{ext}}(i) \Big]. \tag{1.82}$$

The ground-state wave function $|\Psi_\lambda\rangle$ for λ-dependent system is defined so that it satisfies the Schrödinger equation:

$$\hat{H}(\lambda)|\Psi_\lambda\rangle = E(\lambda)|\Psi_\lambda\rangle. \tag{1.83}$$

Then the ground-state total energy, $E(\lambda = 1)$, can be obtained by using Hellmann-Feynman theorem with normalization condition $\langle\Psi_\lambda|\Psi_\lambda\rangle = 1$:

$$E(\lambda = 1) = E_0 + \int_0^1 \mathrm{d}\lambda \langle\Psi_\lambda| \left(\hat{H}_1(\lambda) + \lambda \frac{\mathrm{d}\hat{H}_1(\lambda)}{\mathrm{d}\lambda} \right) |\Psi_\lambda\rangle, \tag{1.84}$$

where $E_0 = \langle\Psi_0|\hat{H}_0|\Psi_0\rangle$ is the zeroth-order energy.

We note that the AC path in DFT is chosen such that the electron density is kept fixed at its physical value along the way, see Ref. [162] for details. The equation 1.84 can be written as follows:

$$\begin{aligned} E &= E_0 + \int_0^1 \mathrm{d}\lambda \langle\Psi_\lambda| \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|r_i - r_j|} |\Psi_\lambda\rangle \\ &\quad + \int_0^1 \mathrm{d}\lambda \langle\Psi_\lambda| \sum_{i=1}^{N} \frac{\mathrm{d}}{\mathrm{d}\lambda} V_\lambda^{\mathrm{ext}}(r_i) |\Psi_\lambda\rangle \qquad (1.85) \\ &= E_0 + \frac{1}{2}\int_0^1 \mathrm{d}\lambda \int\int \mathrm{d}r\,\mathrm{d}r' \langle\Psi_\lambda| \frac{\hat{\rho}(r)[\hat{\rho}(r') - \delta(r-r')]}{|r-r'|} |\Psi_\lambda\rangle \\ &\quad + \int \mathrm{d}r\, \rho(r)[V_{\lambda=1}^{\mathrm{ext}}(r) - V_{\lambda=0}^{\mathrm{ext}}(r)]. \qquad (1.86) \end{aligned}$$

In above equation, $\hat{\rho}(r) = \sum_{i=1}^{N} \delta(r - r_i)$ is electron density operator, and $\rho(r) = \langle\Psi_\lambda|\hat{\rho}(r)|\Psi_\lambda\rangle$ for any $0 \leq \lambda \leq 1$.

The zeroth-order energy for the KS reference state $|\Psi_0\rangle$ given by Slater determinant of occupied single-particle KS orbital $\Psi_i(r)$ can be expressed as follows:

$$E_0 = \langle\Psi_0| \sum_{i=1}^{N} \Big[-\frac{1}{2}\nabla_i^2 + V_{\lambda=0}^{\mathrm{ext}}(r_i)\Big] |\Psi_0\rangle = T_{\mathrm{s}}[\Psi_i(r)] + \int dr \rho(r) V_{\lambda=0}^{\mathrm{ext}}(r). \qquad (1.87)$$

Substituting equation 1.87 into equation 1.85 we get:

$$\begin{aligned} E &= T_{\mathrm{s}}[\Psi_i(r)] + \int \mathrm{d}r \rho(r) V_{\lambda=1}^{\mathrm{ext}}(r) \\ &\quad + \frac{1}{2}\int_0^1 \mathrm{d}_\lambda \int\int \mathrm{d}r\mathrm{d}r' \langle\Psi_\lambda| \frac{\hat{\rho}(r)[\hat{\rho}(r') - \delta(r-r')]}{|r-r'|} |\Psi_\lambda\rangle \qquad (1.88) \\ &= T_{\mathrm{s}}[\Psi_i(r)] + E_{\mathrm{ext}}[\rho(r)] + J[\rho(r)] + \frac{1}{2}\int \mathrm{d}\lambda \int\int \mathrm{d}r\mathrm{d}r' \frac{\rho_{\mathrm{xc}}^\lambda(r,r')\rho(r)}{|r-r'|}. \qquad (1.89) \end{aligned}$$

The last term in above equation is XC energy, in which $\rho_{xc}^\lambda(r,r')$ is XC hole related to density-density correlation function.

$$\rho_{\mathrm{xc}}^\lambda(r,r') = \frac{\langle\Psi_\lambda|\delta\hat{\rho}(r)\delta\hat{\rho}(r')|\Psi_\lambda\rangle}{\rho(r)} - \delta(r-r'), \qquad (1.90)$$

where $\delta\hat{\rho}(r) = \hat{\rho}(r) - \rho(r)$ is "fluctuation" of density operator around its expectation value. This term is linked to response properties (dissipation) of the system through zero-temperature "fluctuation-dissipation" theorem [163]. This theorem then leads to

$$\langle\Psi_\lambda|\delta\hat{\rho}(r)\delta\hat{\rho}(r')|\Psi_\lambda\rangle = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \mathrm{Im}\chi^\lambda(r,r',\omega), \qquad (1.91)$$

where $\chi^\lambda(r,r',\omega)$ is the linear density-response function of the system.

The XC energy can be rewritten with using the term $v(r,r') = \frac{1}{|r-r'|}$ and equation 1.91

$$E_{\mathrm{xc}} = \frac{1}{2}\int_0^1 \mathrm{d}\lambda \int\int \mathrm{d}r\mathrm{d}r' v(r,r')\Big[-\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\omega \mathrm{Im}\chi^\lambda(r,r',\omega) - \delta(r-r')\rho(r)\Big] \tag{1.92}$$

$$= \frac{1}{2}\int_0^1 \mathrm{d}\lambda \int\int \mathrm{d}r\mathrm{d}r' v(r,r')\Big[-\frac{1}{\pi}\int_0^{\infty}\mathrm{d}\omega \chi^\lambda(r,r',i\omega) - \delta(r-r')\rho(r)\Big]. \tag{1.93}$$

Therefore, the task of calculating the XC energy is shifted to that of computing the response functions of a series of fictitious systems along the AC path. The random phase approximation is a particularly simple approximation of the response function:

$$\begin{aligned}\chi^\lambda_{\mathrm{RPA}}(r,r',i\omega) = {}& \chi^0(r,r',i\omega) \\ & + \int \mathrm{d}r_1\mathrm{d}r_2\, \chi^0(r,r_1,i\omega)\lambda v(r_1,r_2)\chi^\lambda_{\mathrm{RPA}}(r_2,r',i\omega),\end{aligned} \tag{1.94}$$

where $\chi^0(r,r_1,i\omega)$ is the independent-particle response function of the KS reference system at $\lambda = 0$.

$$\chi^0(r,r',i\omega) = \sum_{i,j}\frac{(f_i - f_j)\Psi_i^*(r)\Psi_j(r)\Psi_j^*(r')\Psi_i(r')}{\epsilon_i - \epsilon_j - i\omega}. \tag{1.95}$$

where ϵ_i are orbital energies and f_i are occupation factors.

Finally, the XC energy in RPA can be separated into an exact exchange (EXX) and the RPA correlation term.

$$E_{\mathrm{x}}^{\mathrm{EXX}} = \frac{1}{2}\int\int \mathrm{d}r\mathrm{d}r' v(r,r')\Big[-\frac{1}{\pi}\int_0^{\infty}\mathrm{d}\omega \chi^0(r,r',i\omega) - \delta(r-r')\rho(r)\Big] \tag{1.96}$$

$$= -\sum_{ij} f_i f_j \int\int \mathrm{d}r\mathrm{d}r' \Psi_i^*(r)\Psi_j(r)v(r,r')\Psi_j^*(r')\Psi_i(r') \tag{1.97}$$

and

$$E_{\mathrm{c}}^{\mathrm{RPA}} = -\frac{1}{2\pi}\int\int \mathrm{d}r\mathrm{d}r' v(r,r')\int_0^{\infty}\mathrm{d}\omega\Big[\int_0^1 \mathrm{d}\lambda \chi^\lambda_{\mathrm{RPA}}(r,r',i\omega) - \chi^0(r,r',i\omega)\Big] \tag{1.98}$$

$$= \frac{1}{2\pi}\int_0^{\infty}\mathrm{d}\omega \mathrm{Tr}[\ln(1-\chi^0(i\omega)v) + \chi^0(i\omega)v]. \tag{1.99}$$

The RPA method offers several advantages over standard DFT approximations and other methods. The RPA correlation energy is inherently non-local, making it well-suited for capturing long-range interactions, such as van der Waals forces. RPA incorporates the exact-exchange energy, which cancels exactly the self-interaction error present in the Hartree energy. Moreover, RPA considers dynamic electronic screening effects by summing up a sequence of "ring" diagrams to infinite order. This makes RPA particularly effective for small-gap systems, where finite-order MBPT methods often fail [42, 164].

1.5.2 Beyond-RPA methods

The RPA method has been known to underestimate binding energies in many cases [41, 43, 165, 166]. The accuracy of RPA can be improved by including additional terms, such as higher-order exchange terms or various singles corrections, or by including approximate exchange–correlation kernels [41, 167, 168, 169, 170, 171, 172, 173, 174]. Moreover, most current RPA calculations are performed non-self-consistently, using DFT orbitals as input [160, 175]. There have been several studies that analyzed whether performing self-consistent RPA calculations affects the accuracy [176, 177, 178, 179, 180] However, these methods can significantly increase computational demands and are not yet implemented in the PBC approach. In this section, we will discuss the beyond-RPA contributions (corrections) we used in our work.

The total RPA energy without any corrections can be expressed as follows:

$$E_{\mathrm{tot}}^{\mathrm{RPA}} = E^{\mathrm{EXX}} + E_{\mathrm{c}}^{\mathrm{RPA}}. \tag{1.100}$$

In above equation, E^{EXX} is the HF energy evaluated with approximate DFT orbitals and $E_{\mathrm{c}}^{\mathrm{RPA}}$ is direct RPA correlation energy.

The common correction is the inclusion of singles contributions, which arise because Brillouin's theorem does not apply when using DFT orbitals for RPA calculations. This correction can be calculated by using mean-field 1-RDMs built from the eigenvectors of the Fock matrix $h[\gamma^0]$

$$\Delta E^{\mathrm{HF}} = E^{\mathrm{HF}}[\gamma] - E^{\mathrm{HF}}[\gamma^0], \tag{1.101}$$

where γ is the fully interacting 1-RDM and γ^0 is the noninteracting density matrix.

The total singles correction can be divided into linear contribution (E^{RSE}) and the quadratic remainder ($E^{\mathrm{1RDM,quad}}$):

$$E^{\mathrm{RSE}} = \sum_{ij}^{orb} \delta\gamma_{ij}(h[\gamma^0])_{ij}. \tag{1.102}$$

$$E^{\mathrm{1RDM,quad}} = \frac{1}{2}\sum_{ijkl}^{orb}\Big(\delta\gamma_{ij}\delta\gamma_{kl} - \frac{1}{2}\delta\gamma_{il}\delta\gamma_{kj}\Big)(ij|kl). \tag{1.103}$$

In these equations, i, j, k, l are general (spin) orbitals, $\delta\gamma$ represents the difference between the fully interacting and non-interacting 1-RDMs, $\delta\gamma = \gamma - \gamma^0$. The RSE contribution was derived in Refs. [41, 181]. The 1RDM,quad contribution is a straightforward extension of RSE. Detailed derivation of this energy component can be found in Ref. [45]. Note that the singles energy contributions vanish for the self-consistent Hartree-Fock reference due to Brillouin theorem.

As mentioned before, the RPA correlation energy can be derived from an approximate coupled-cluster doubles theory. This derivation was shown by Scuseria *et al.* [156]

$$E_{\mathrm{c}}^{\mathrm{RPA}} = 2\sum_{rasb}^{orb}(ra|sb)T_{ab}^{rs}. \tag{1.104}$$

In this context, the doubles amplitudes, T_{ab}^{rs}, are obtained by solving the RPA doubles amplitude equation. These amplitudes only consider the ring (RPA)

term and do not include any exchange interactions. Therefore, the second-order screened exchange correction (SOSEX) offers an alternative approach to extend beyond the direct RPA [182]. This correction was derived from couple cluster theory [156], and removes a subset of the Pauli-exclusion violating contributions from RPA correlation energy

$$E_{\mathrm{c}}^{\mathrm{SOSEX}} = -\sum_{rasb}^{orb}(sa|rb)T_{ba}^{sr}. \tag{1.105}$$

The SOSEX correction exhibits rapid decay with distance between molecules in interactions. Therefore, this correction is expected to affect mostly compact clusters. This leads to significant improvement the short-range over-correlation probem in RPA [160, 183].

The simplifying assumptions of direct RPA correlation energy also neglect third- and higher-order many-body perturbation theory terms, which involve interactions between hole-hole and particle-particle orbital pairs. In our study, we include the third-order correction, 2g, which was derived in detail in Ref. [45]

$$E_{\mathrm{c}}^{\mathrm{2g}} = -4\sum_{rasbtc}(ab|rs)T_{bc}^{rt}T_{ca}^{ts}. \tag{1.106}$$

The importance of 2g correction to RPA has been observed particularly in noncovalent systems [45, 183]. This correction decays slowly, proportional to $1/R^6$, with R is distance between molecules.

The addition of the RPA corrections discussed above to direct RPA energy leads to three variants of beyond-RPA methods we used to examine in paper P2.

The basic variant is RPA+RSE approach [41, 166]

$$E_{\mathrm{tot}}^{\mathrm{RPA+RSE}} = E^{\mathrm{EXX}} + E_{\mathrm{c}}^{\mathrm{RPA}} + E^{\mathrm{RSE}}. \tag{1.107}$$

The higher-level variant is renormalized second-order perturbation theory (rPT2) approach [181]

$$E_{\mathrm{tot}}^{\mathrm{rPT2}} = E^{\mathrm{HF}} + E_{\mathrm{c}}^{\mathrm{RPA}} + E^{\mathrm{RSE}} + E_{\mathrm{c}}^{\mathrm{SOSEX}}. \tag{1.108}$$

Finally, the variant, which uses all corrections specified is referred to as RPA+MBPT3 [45]

$$\begin{aligned} E_{\mathrm{tot}}^{\mathrm{RPA+MBPT3}} = {} & E^{\mathrm{HF}} + E_{\mathrm{c}}^{\mathrm{RPA}} + E^{\mathrm{RSE}} \\ & + E^{\mathrm{1RDM,quad}} + E_{\mathrm{c}}^{\mathrm{SOSEX}} + E_{\mathrm{c}}^{\mathrm{2g}}. \end{aligned} \tag{1.109}$$

1.6 Basis sets

As mentioned before, it is necessary to use a basis-set expansion to express the unknown spatial orbitals in terms of well-defined functions for practical calculations in quantum mechanics. The choice of an appropriate basis set depends on specific properties of the considered systems and desired level of precision. There are two main types of basis sets commonly used. Specifically, localized basis sets are typically employed in molecular and cluster calculations, whereas plane-wave basis sets are often used in solid-state calculations. We use both of them for

our calculations, and therefore we will discuss them in this section, starting with localized basis sets.

Localized basis sets consist of atomic or molecular orbitals that are centered on specific atoms or groups of atoms. The Slater-type functions introduced by Slater [184] seem to be the most natural choice of the atomic orbitals. This is because the radial part of Slater orbitals decays in the same way as the exact solutions of Schrödinger equation for hydrogen atom. However, the evaluation of molecular integrals with Slater-type orbitals can not be done analytically. To overcome this difficulty, Francis S. Boys [185, 186] recommended the use of Gaussian orbitals, which lead to simpler integral evaluation thanks to Gaussian product theorem. On the other side, the radial shape of the Gaussian orbitals is not as good as one of the Slater orbitals as they have too fast decay with increasing distance from nucleus and no cusp at very small distances. Therefore, it is necessary to use more Gaussian than Slater basis functions in the calculation to achieve the same precision. To capture the correct radial behavior of Slater-type orbitals, a linear combination of multiple Gaussian functions is used to represent one basis function. This type of a basis function is called a contraction.

The size of a localized basis set is a crucial factor. For isolated atoms, a minimal basis set, such as a single $1s$ orbital for a hydrogen atom, may be sufficient to provide a reasonable approximation for wavefunction. However, when atoms form chemical bonds with other atoms, this minimal basis set is often insufficient to describe the shape of wavefunction, and it is necessary to use larger basis sets. There are several standard basis sets which have been proposed to improve the minimal basis set. One of them is Pople basis sets [187, 188], which have been widely used in quantum chemistry for many years, especially for HF and DFT calculations. However, in order to recover a significant part of the correlation energy, correlation-consistent basis sets introduced by Dunning and coworkers, are widely used [189, 190, 191]. These basis sets are often denoted as cc-pVXZ, where cc represents for correlation-consistent, p for polarized, V for valence, the number X is the cardinal number (X = D, T, Q, 5, and so on), and Z is Zeta. They can be then systematically improvable by adding diffuse functions to capture electron correlation more effectively. In this way, one obtains aug-cc-pVXZ basis sets, which we use in our MBE calculations. The main advantage of Dunning's correlation-consistent basis sets is that they are systematically constructed with increasing completeness, which is suitable for extrapolating to complete basis set (CBS) limit.

The use of a small basis sets offers computational efficiency, yet it may result in less precision. Furthermore, using finite basis sets can introduce the basis set superposition error (BSSE). This error arises from the fact that when two or more molecules come together to form a complex, the basis functions of one molecule can influence the optimization of orbitals on the other molecule. This can lower the total energy of the system and lead to an overestimation of the binding energy. To address this error, we use large basis sets or extrapolate results to CBS depending on considered energy components. Alternatively, counterpoise correction methods can be employed [192, 193, 194]. Specifically, we evaluate energies of monomers in dimer basis and similar.

The basis sets we use for our calculations in the PBC approach are plane wave basis sets. These basis sets are a popular choice for crystal calculations due

to their suitability for modeling these systems. The orbitals describing electron distribution in a crystal can be expanded using a set of plane wave functions e^{iGr}, where G is a vector of the reciprocal lattice. Here, Fourier representation is used to describe how electronic wave functions are expressed in terms of plane waves. In this representation, the coefficients of plane wave expansion are Fourier coefficients. By computing and storing these coefficients, we can describe electronic structure of solid and other properties of interest. The size of the basis set is controlled by the highest momentum of plane wave vectors that are included in expansion. The plane wave basis sets can be used in combination with pseudopotentials, which simplify the treatment of all electrons in an atom by replacing core electrons and their interactions with nucleus with an effective potential that acts on valence electrons. A more advanced treatment of the core electrons is provided by the projector-augmented wave (PAW) method which is implemented in the Vienna ab-initio simulations package (VASP) that we used for PBC calculations [195, 196]. The plane wave basis sets offer several advantages. One notable benefit is their orthogonality, which simplifies the evaluation of various terms in electronic Hamiltonian. Moreover, they describe entire crystal cell uniformly, thus there is no BSSE error. Despite their advantages, it is important to note that they are very inefficient when studying isolated molecules. This is because the same quality of basis set is used in the regions where there is no electron density. Furthermore, the evaluation of Fock exchange term with plane wave basis sets is more computationally expensive in comparison with localized basis sets.

1.7 Computational techniques

In this section, we will first focus on discussing periodic systems (molecular solids). We then discuss how the binding energy of these systems can be determined through periodic boundary conditions and many-body expansion approaches. The advantages and limitations of these approaches are also discussed in detail. Finally, we will introduce the correction scheme used in our work.

1.7.1 Periodic systems

Molecular solids are periodic systems, containing an infinite number of electrons and nuclei. Thus, we have to find a suitable way to characterize such systems. In molecular solids, atoms or molecules are repeated in space. This regular repetition is determined by lattice vectors $\vec{R}$, which can be formed as a linear combination of so-called unit-cell vectors $\vec{a_i}$ as follows:

$$\vec{R} = n_1\vec{a_1} + n_2\vec{a_2} + n_3\vec{a_3} = \sum_{i=1}^{3} n_i\vec{a_i}. \tag{1.110}$$

Similarly, we may construct reciprocal lattice of vectors $\vec{G}$ as:

$$\vec{G} = n_1\vec{b_1} + n_2\vec{b_2} + n_3\vec{b_3} = \sum_{i=1}^{3} n_i\vec{b_i}, \tag{1.111}$$

where $\vec{b_i}$ are primitive reciprocal lattice vectors that are defined as:

$$\vec{a_i} \cdot \vec{b_i} = 2\pi\delta_{ij}. \tag{1.112}$$

We note that $\vec{a_i}$ vectors are not unique, and we use a specific primitive cell known as Wigner-Seitz cell. This cell is related to first Brillouin zone (BZ) in reciprocal space. The volume of the first BZ (V_{BZ}) is then defined as:

$$V_{BZ} = \frac{(2\pi)^3}{V_0}, \tag{1.113}$$

where $V_0 = \vec{a_1} \cdot (\vec{a_2} \times \vec{a_3})$ is the volume of the unit cell.

In the context of an infinite lattice, it is not possible to normalize a wave function over all space because the lattice extends infinitely. Therefore, we introduce a normalization volume $\Omega = l^3 V_0$, in which l is an integer much greater than 1 to ensure that Ω is large enough to effectively model the infinite system. All points that are outside Ω are then mapped onto a point inside Ω through periodic boundary conditions.

We then introduce a translation operator acting on the wavefunction that changes it by nothing more than a phase factor

$$\hat{T}_R \Psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\vec{R}} \Psi_{\vec{k}}(\vec{r}). \tag{1.114}$$

Here we restrict our considerations to single-particle eigenfunctions for electrons. $\vec{k}$ is a vector and may be considered as a "quantum number" characterizing the wave functions of a periodic crystal. The eigenfunctions and eigenvalues can then be obtained by considering only $\vec{k}$ in the first BZ. This means that we can solve the equations for a set of k-points instead of simulating a large supercell.

The eigenfunctions are given by the Bloch's theorem, which states that any single-particle wave function in a periodic system can be written as the product of a plane wave and a lattice-periodic part:

$$\Psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\vec{r}} \phi_{\vec{k}}(\vec{r}). \tag{1.115}$$

The energy eigenvalues are periodic in reciprocal space:

$$\epsilon(\vec{k}) = \epsilon(\vec{k} + \vec{G}). \tag{1.116}$$

In principle, we should know eigenvalues and eigenfunctions at all k-points in first BZ. However, in practice, the wave functions at k-points that are close together in BZ are almost identical due to periodicity of crystal lattice. This enables approximation of integration over $\vec{k}$ by a weighted sum over a discrete set of points, rather than performing a continuous integration over entire BZ. Numerous methods have been suggested for sampling k-points in this context. Among these, Monkhorst-Pack [197] is one of the most widely used scheme.

For periodic systems, to make the calculations practical, the plane waves are introduced as basis functions for the expansion of unknown Bloch waves as follows:

$$\Psi_{\vec{k},j}(\vec{r}) = \sum_{\vec{G}} C_{\vec{G},\vec{k},j} \frac{1}{\sqrt{\Omega}} e^{i(\vec{k}+\vec{G})\cdot\vec{r}}. \tag{1.117}$$

In above equation, j is energy band index. The sum over reciprocal lattice vectors $\vec{G}$ is limited to those for which $\frac{\hbar^2}{2m}(\vec{k}+\vec{G})^2 \leq E_{\text{cutoff}}$. In other words, only plane waves with kinetic energy below the specified cutoff are included in basis set used for the calculations.

1.7.2 Approaches for calculations of binding energy of molecular solids

Binding energy of a crystal with PBC approach can be obtained by this expression:

$$E_b = \frac{E_{\text{solid}}}{Z} - E_{\text{mol}}, \tag{1.118}$$

where E_{solid} is energy of a crystal, E_{mol} is energy of an isolated molecule, and Z is number of molecules in unit cell. In this approach, the binding energy is influenced by various numerical parameters, including possible approximations of core electrons, energy cutoff of plane-wave basis set, density of k-points used for solid calculations, and volume of unit cell used for molecule calculations. Obtaining converged values with these parameters can significantly increase the computational time [16, 21, 23]. This has currently limited the use of PBC approach with reference methods, such as CCSD(T) [29, 198, 199].

As an alternative, binding energy of molecular solids can be determined using MBE calculations. The basic idea of this approach is to decompose the energy (or other properties) of a large system into a sum of smaller contributions. This reduces the computational requirements compared to the evaluation of single energy in the PBC approach. Therefore, CCSD(T) method becomes feasible to perform in the MBE approach, but still only for small systems.

In this approach, the energy of a system with N monomers can be expanded in orders of interaction energies as follows:

$$E = \sum_{i=1}^{N} E_i + \sum_{i<j}^{N} \Delta E_{i,j} + \sum_{i<j<k}^{N} \Delta E_{i,j,k} + \sum_{i<j<k<l}^{N} \Delta E_{i,j,k,l} + \dots \ , \tag{1.119}$$

where $\Delta E_{i,j}$ is the interaction energy of dimers, and $\Delta E_{i,j,k}$ and $\Delta E_{i,j,k,l}$ are non-additive energies of trimers and tetramers, respectively. They can be expressed as follows:

$$\Delta E_{i,j} = E_{i,j} - E_i - E_j. \tag{1.120}$$

$$\Delta E_{i,j,k} = E_{i,j,k} - \Delta E_{i,j} - \Delta E_{i,k} - \Delta E_{j,k} - E_i - E_j - E_k. \tag{1.121}$$

$$\begin{aligned} \Delta E_{i,j,k,l} = E_{i,j,k,l} &- \Delta E_{i,j,k} - \Delta E_{i,j,l} - \Delta E_{i,k,l} - \Delta E_{j,k,l} \\ &- \Delta E_{i,j} - \Delta E_{i,k} - \Delta E_{i,l} - \Delta E_{j,k} - \Delta E_{j,l} - \Delta E_{k,l} \\ &- E_i - E_j - E_k - E_l, \end{aligned} \tag{1.122}$$

In above equations, E_i is the energy of monomer i. $E_{i,j}$ is the energy of dimer, and $E_{i,j,k}$, and $E_{i,j,k,l}$ are the energies of trimer and tetramer, respectively.

The higher-order terms can be defined similarly.

In the case of molecular solids where all molecules are symmetry equivalent, we select one of the molecules as a reference (ref). The binding energy of a molecular solid, $\mathrm{E_b}$, is then obtained from interaction energies of dimers and non-additive three-, four-, and higher-body energies as:

$$E_b = \frac{1}{2}\sum_{j} \Delta E_{ref,j} + \frac{1}{3}\sum_{j<k} \Delta E_{ref,j,k} + \frac{1}{4}\sum_{j<k<l} \Delta E_{ref,j,k,l} + \dots \ . \tag{1.123}$$

In our work, we assume that the gas phase structure is identical to the solid phase structure, resulting in the monomer term being zero in equation 1.123. In the MBE approach, there are several numerical parameters we need to consider to reach converged results, including the choice of fragment size (dimer, trimer,...), the number of fragments included for each specific size, and the basis-set size [17, 28, 30, 200, 201, 202]. The overall computational expense is significantly dependent on how rapidly MBE achieves convergence with these parameters. Moreover, the MBE convergence depends on the considered systems. Specifically, the convergence with the fragment size can be slow for the systems where many-body polarization effects are important [202, 203]. This issue can be addressed by using techniques such as fragment embedding [24, 204, 205] or by using a force field to account for more distant fragments [206]. However, the precision and accuracy of these approaches remains a subject of uncertainty [207].

In principle, all individual contributions in MBE should be included to obtain the total energy with a high precision. However, in practice, it is common to neglect high-order terms (four-body and higher-order terms) in the expansion to make the calculations more tractable because these terms are expected to have small contributions [23, 208]. Moreover, the precise evaluation of high-order terms can be difficult due to numerical errors [202]. However, high-order terms have been shown to be important in predicting relative conformational energies of proteins [209], and in systems where many-body polarization effects are significant [202, 203, 210]. Therefore, the decision to include or neglect the high-order terms in MBE is dependent on the considered systems and the level of desired precision.

In equation 1.123, the summations are infinite as they run over all molecules in crystal. However, in practice, cut-off distances are introduced. This is because the interaction energy of individual contributions at large distances is very small and can be usually neglected [23]. In our work, we defined distance of dimers as average Cartesian distance of all pairs of atoms between two molecules. For trimers and tetramers, the distance is determined by summing the distances of all the dimers contained within the cluster.

1.7.3 Correction scheme

The CCSD(T) method can be applied in the MBE approach to obtain reliable binding energy of molecular solids, but it is typically limited to small systems. As discussed earlier, apart from the basis set size, it is necessary to calculate each individual contribution in MBE, ensuring convergence with the number of molecules in each contribution (cut-off distance). This makes reaching the reliable binding energy of large molecular solids with CCSD(T) method more challenging. Therefore, the development of a scheme to address this difficulty is important.

It is expected that the difference in binding energy between CCSD(T) method and approximate methods occurs mainly at short distances. This is because interactions between molecules typically decrease as the distance between them increases. Based on this fact, we tried to examine the accuracy of the correction scheme where a portion of the CCSD(T) data is replaced with results obtained from approximate methods [211]. Clearly, this scheme can be particularly beneficial for a substantial reduction of the computational cost. The binding energy

in the correction scheme can be expressed as follows:

$$E_b^{\mathrm{A}} = E_b^{\mathrm{B}} + \frac{1}{2}\sum_{j}^{\%}\left(\Delta E_{ref,j}^{\mathrm{A}} - \Delta E_{ref,j}^{\mathrm{B}}\right) + \frac{1}{3}\sum_{j<k}^{\%}\left(\Delta E_{ref,j,k}^{\mathrm{A}} - \Delta E_{ref,j,k}^{\mathrm{B}}\right) + \frac{1}{4}\sum_{j<k<l}^{\%}\left(\Delta E_{ref,j,k,l}^{\mathrm{A}} - \Delta E_{ref,j,k,l}^{\mathrm{B}}\right). \tag{1.124}$$

In above equation, % is percentage of the fragments used for replacement, A and B represent reference CCSD(T) and approximate methods, respectively. In our work we tested how suitable different methods are for the correction scheme and how large is the percentage of clusters that needs to be treated explicitly.

1.8 Methods for calculations of binding energy of molecular solids

The calculations of reliable binding energy of molecular solids, converged with numerical parameters using CCSD(T) method, are highly demanding and impractical for large systems. In this section, we will provide an overview of the approximate methods currently used to calculate the binding energy of molecular solids.

In principle, empirical force fields can be applied to calculate binding energy of molecular solids [212, 213, 214]. Their notable advantage lies in their low computational expense, making them suitable for calculating the properties of large systems. However, they may not capture the subtle balances between intra- and intermolecular interactions that occur in molecular crystals. Additionally, the quality of predictions heavily depends on the quality of the force field parameters, which may not always be readily available for all the systems. Furthermore, the available force fields typically do not account for polarization. Nevertheless, recent advancements in machine learning force fields (ML-FF) can yield binding energies that are comparable with the results obtained with quantum mechanical methods [26, 215]. For instance, the performance of the ML-FF approach proposed by Thürlemann *et al.* is comparable to the most accurate dispersion-corrected DFT method reported so far for crystals in X23 dataset and 13 ice phases [215].

Over the past decade, electronic structure methods have been used widely to describe interactions in molecular solids. Among them, semi-empirical methods, which can be derived from either HF or DFT theory by introducing empirical parameters and methodological approximations, are the simplest variants [216]. These methods, especially density-functional tight-binding, have emerged as promising tools to calculate binding energies of molecular solids due to their computational efficiency [217, 218, 219]. However, there are still several limitations that hinder the use of these methods as the standard way for predicting the stability of molecular crystals [220, 221].

DFT methods can be used to calculate binding energies of molecular solids. However, the performance of these methods is influenced by the choice of DFT functionals and the specific types of molecular solids studied [71]. For example,

standard DFT approximations can not describe exactly charge density and dispersion interactions. The use of hybrid DFT functionals can mitigate delocalization errors and offer higher accuracy [37, 222]. Additionally, dispersion-corrected DFT models have been developed and used to obtain binding energies of molecular solids with accuracy sufficient for many purposes [33, 34, 35, 36, 146, 152, 223, 224]. As an illustration, employing the DFT method with a new XDM correction results in a mean absolute error of merely 0.8 kJ/mol for lattice energies of 13 ice phases, surpassing the performance of all previously reported DFT functionals [36]. Nevertheless, it is important to note that these DFT models may not always provide high accuracy [21, 22, 63, 225]. For example, several dispersion-corrected DFT models have shown to underestimate the binding energy of CO_2 crystal by more than 10 % [21]. Double hybrid DFT methods have also demonstrated to provide good accuracy for binding energies of molecular solids [226, 227, 228, 229]. However, their performance was shown not to be superior to wave function methods, such as MP2 [228].

The alternative to DFT methods for obtaining binding energies of molecular solids are wave function methods based on perturbation theory, such as Møller–Plesset perturbation theory or random phase approximation. These methods do not reach the accuracy of coupled cluster [16] but provide more consistent and offer more accurate results compared to standard DFT approximations [21]. Among such methods, MP2 theory is the simplest form of many-body perturbation theory and is sometimes used [40, 230, 231, 232, 233, 234]. This method can yield binding energies which are in a very good agreement with the reference CCSD(T) values for some crystals, such as CH_4 and CO_2 [23]. However, the quality of binding energies obtained from MP2 method is affected by the poor performance in systems with delocalized electrons [235], and the lack of three-body and higher dispersion terms [236]. Recently, spin-component-scaled (SCS) MP2 variant, which scales the same-spin and opposite-spin components of the correlation energy, has demonstrated superior performance over standard MP2 in describing the binding energies of crystals [234, 237].

RPA method can overcome the limitations of MP methods thanks to satisfactory description of various bonding interactions. This makes the RPA method suitable for describing the binding energies of molecular solids [16, 21, 238, 239]. However, it has known that standard RPA method tends to underestimate the binding energies due to its reliance on non-self-consistent DFT orbitals [21, 41, 240]. For instance, the RPA binding energies of studied molecular solids underestimate the reference data by more than 10 % [21]. Currently, single corrections are added to RPA to improve the accuracy of RPA for binding energies of crystals [16, 21]. In this thesis we focus on RPA and try to understand where its errors come from and how they can be reduced.

1.9 Summary

In this chapter, we described extensively the theoretical methods, including HF, DFT, RPA, MP, and CC, currently used to calculate the properties of materials, especially for molecules and solids. Additionally, we discussed the methodologies for determining the binding energies in molecular solids using both PBC

and MBE approaches. In the following we use the methods with a goal of understanding their accuracy and the effect of numerical parameters on the results.

2. Computational details

To perform the calculations and generate the results for analysis in this thesis, we first selected a set of molecular solids. Subsequently, we carried out the calculations of the studied methods using both MBE and PBC approaches to obtain binding energies of the selected systems. Below, we provide a detailed description of the computational settings used for these calculations.

2.1 Systems

Four different molecular solids were selected, including monoclinic ethane [241] and ethylene [242], and orthorhombic and cubic forms of acetylene [243]. To differentiate between the two acetylene forms, they are denoted as "acetylene/I" for the orthorhombic form and "acetylene/II" for the cubic form from this point onward. We note that in our papers (P1 and P2 in List of Publications), we used "acetylene/c" and "acetylene/o" for the orthorhombic and cubic forms, respectively, but we changed the notation in this thesis. The selection of these systems is motivated by several key considerations. Firstly, they are held together by dispersion forces, which are challenging to describe accurately by electronic structure methods. Furthermore, there is a variation in the importance of electrostatic contributions from ethane to acetylene. This can have an effect on the convergence of the energy components with the numerical parameters. Lastly, they are compact in size, making them possible for reference CCSD(T) calculations.

Table 2.1: The unit cell volume at experimental equilibrium (V_0), the lattice constants (a, b, c), the number of molecules in the unit cell (Z), and the CSD code of the selected molecular solids.

Systems	V_0(Å)	Lattice constants			Z	CSD code
		a	b	c		
Ethane	138.89	4.226	5.623	5.845	2	ETHANE01
Ethylene	124.18	4.626	6.620	4.067	2	ETHLEN01
Acetylene/I	208.23	6.198	6.023	5.578	4	ACETYL11
Acetylene/II	227.54	6.105	6.105	6.105	4	ACETYL03

The crystal structures of the selected systems were obtained from the Cambridge Structural Database (CSD) [244], as detailed in Table 2.1. Using the optB88-vdW functional [245, 246, 247, 248], we optimized atom positions of the selected systems while keeping the lattice parameters at their experimental values. Subsequently, we directly extracted the geometries of isolated molecules from these optimized crystal structures and used them without additional optimization for the MBE calculations using mbe.py library in Ref. [23].

2.2 Orbitals used for RPA calculations

Within the PBC approach, four KS-DFT orbitals were selected: KS/PBE, KS/SCAN, KS/PBE0, and KS/SCAN0. Apart from these orbitals, we used semicanonical/PBE (abbreviated as semi/PBE) and HF orbitals within the MBE approach. Here the semi/PBE orbitals are eigenstates of the occupied-occupied and virtual-virtual blocks of the Fock matrix computed from KS/PBE occupied orbitals [249, 250, 251]. This means that they serve as intermediate choice between KS/PBE and HF orbitals. For the RPA corrections in the rPT2 and RPA+MBPT3 methods, we used only three types of orbitals, including KS/PBE, semi/PBE, and HF orbitals. The only exception here are the RSE corrections, for which all examined orbitals were used.

2.3 Computational techniques

We begin by describing the techniques in the MBE approach. To generate structures of the fragments (dimers, trimers, ...) for the calculations, we used an in-house library from Ref. [23]. The identification of symmetry equivalent clusters was carried out using the approach proposed in Ref. [252]. The Molpro program [253] was used for the MP2, MP3, and CCSD(T) calculations. For the RPA and its additional corrections, an in-house code using a canonical-orbital variant of the algorithm described in Ref. [43] was used. All the correlation energies were obtained within the frozen-core approximation.

The MBE calculations were first calculated up to the 4-body term for the CCSD(T) calculations. We then found that for our systems the contribution of the 4-body CCSD(T) energy is very small, and its accurate evaluation can be challenging due to numerical errors as observed before [202]. Therefore, we decided to truncate the MBE calculations at the 4-body term for MP2 and RPA based on pure DFT orbitals. For MP3 and RPA based on semi/PBE or HF orbitals, the calculations were limited to the 3-body term. Similarly, all RPA corrections were determined up to the 3-body term, with the exception of RSE, which was extended up to the 4-body term. The detailed discussion of the used settings follows in the next chapter.

Dunning's augmented correlation-consistent basis sets were used for all the MBE calculations [254]. The extrapolations to CBS limit using the formula of Halkier *et al.* [255] were applied to reduce the basis-set incompleteness errors of the correlation energies.

$$E_{\mathrm{CBS}} = \frac{(X+1)^n E_{X+1} - X^n E_X}{(X+1)^n - X^n}\,, \tag{2.1}$$

where E_X is the energy in the AVXZ basis set, and $n = 3$ was set. The complete auxiliary basis set singles corrections (CABS) [256, 257] to the HF energy were calculated to improve the basis set convergence of the HF component. To mitigate the basis-set dependence of the MP2 and CCSD correlation energies, the F12 methods were used in some cases [258, 259]. In these cases, the extrapolations are made using equation 2.1 with $n = 5$ [23, 260]. For the triples (T) component in the CCSD(T) method, scaling was applied to the two-body term [261], while it

remained unscaled for the three- and four-body contributions. This is because the non-additive (T) contributions converge quickly with the basis set size, resulting in almost no difference between the scaled and unscaled (T) values [44].

We now move to describe the techniques in the PBC approach. We used VASP [195, 196, 262], which implements PAW method [263, 264], to perform the RPA and MP2 calculations. More specifically, the RPA calculations followed the algorithm described in Ref. [265, 266]. The RSE energies were evaluated according to Ref. [41]. For the MP2 calculations, the standard algorithm implemented in VASP by Marsman *et al.* was used [267]. The "standard" PAW data sets based on KS/PBE states were employed for the correlation energies. For the mean-field energies, the "hard" PAW data sets were used, except for the HF and RSE components, which used "standard" PAW data sets. In the second part of the next chapter, we discuss the convergence of the energies with respect to various parameters that are needed to obtain converged binding energies.

3. Convergence of energy components with numerical parameters

The energy components calculated from both MBE and PBC approaches depend on various numerical parameters. It is essential to determine the optimal settings for these parameters that are reliable and computationally feasible for obtaining very precise energy values. In this chapter, we first focus on understanding the challenges associated with reaching converged values of the energy components with respect to the numerical parameters. We then try to estimate the uncertainties in the energy components to make comparison of their results between MBE and PBC approaches more reliable in the next chapter. The convergence of the energy components with numerical parameters in the MBE calculations was discussed in our papers P1 and P2, while the convergence of the energy components in the PBC calculations is discussed in our upcoming work.

3.1 Many body expansion

In the MBE approach, the results of n-body contributions are affected by two main parameters: basis set and cut-off distance. We need to obtain converged values of the energy components with these two parameters for comparison with PBC results in chapter 4 and examination of approximate methods in chapter 6. The effect of these parameters will be discussed in detail in this section, starting with the basis set, then with the cut-off distance. We note that obtaining converged values with the cut-off distance is only necessary when using MBE results to compare to PBC results or when calculating a reference binding energy. For testing of accuracy, it is sufficient to use a non-converged n-body contribution.

3.1.1 Basis set

It is critical to understand and identify which basis set is reliable enough for each energy component, as the computational cost increases rapidly with the basis-set size. To evaluate the convergence of an energy component with the basis-set size, it is typically necessary to compute this energy using various basis sets. However, this can be computationally demanding. Here, we only evaluate the convergence of the energy components with the basis set based on difference between the values obtained with the AVQZ and AVTZ basis sets. We discuss the convergence of each of n-body terms separately, starting with the two-body contributions, then with the non-additive contributions.

3.1.1.1 Two-body terms

The basis set convergence of the energy components obtained from the 2-body calculations was discussed in our papers P1 and P2. Specifically, we focused on the energy components in the MP2 and CCSD(T) methods in paper P1, and

Table 3.1: The basis set convergence of the energy components obtained from the 2-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets. The energy components (E^{DFT}, E^{EXX}, and E^{RSE}) do not depend on the transformation from KS/PBE orbitals to semi/PBE orbitals, and are thus identical.

Inputs	Components	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	E^{DFT}	0.00	−0.03	−0.15	−0.25
	E^{EXX}	−0.04	−0.05	−0.09	−0.07
	E^{RSE}	0.02	0.02	0.05	−0.02
	E_c^{RPA}	−0.63	−0.81	−1.32	−0.98
	E_c^{SOSEX}	0.13	0.18	0.31	0.22
semi/PBE	E^{DFT}	0.00	−0.03	−0.15	−0.25
	E^{EXX}	−0.04	−0.05	−0.09	−0.07
	E^{RSE}	0.02	0.02	0.05	−0.02
	$E^{\mathrm{1RDM,quad}}$	0.01	0.00	0.00	0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.58	−0.73	−1.13	−0.87
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	0.14	0.19	0.31	0.22
	$E_{\mathrm{c}}^{\mathrm{2g}}$	−0.10	−0.15	−0.25	−0.19
KS/SCAN	E^{DFT}	−0.13	−0.09	−0.17	−0.11
	E^{EXX}	0.11	0.15	0.07	0.02
	E^{RSE}	−0.12	−0.15	−0.13	−0.11
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.06	−0.29	−1.04	−0.76
KS/PBE0	E^{DFT}	−0.01	−0.01	−0.09	−0.17
	E^{EXX}	−0.02	−0.01	−0.03	−0.07
	E^{RSE}	0.00	0.00	−0.01	−0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.59	−0.78	−1.27	−0.96
KS/SCAN0	E^{DFT}	−0.04	0.00	−0.07	−0.02
	E^{EXX}	0.04	0.07	0.04	0.00
	E^{RSE}	−0.06	−0.09	−0.09	−0.08
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.22	−0.45	−1.10	−0.82
HF	E^{HF}	−0.02	−0.01	−0.04	−0.07
	$E^{\mathrm{HF+CABS}}$	−0.01	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.71	−0.97	−1.50	−1.18
	$E_{\mathrm{c}}^{\mathrm{MP2}}$	−0.54	−0.74	−1.18	−0.92
	$E_{\mathrm{c}}^{\mathrm{MP2-F12}}$	−0.03	−0.09	−0.60	−0.45
	$E_{\mathrm{c}}^{\mathrm{MP3}}$	−0.50	−0.71	−1.18	−0.94
	$E_{\mathrm{c}}^{\mathrm{CCSD}}$	−0.31	−0.53	−0.95	−0.75
	$E_{\mathrm{c}}^{\mathrm{CCSD-F12b}}$	−0.02	−0.10	−0.40	−0.45
	$E_{\mathrm{c}}^{\mathrm{T}}$(scaled)	0.03	0.00	−0.04	−0.04
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	0.20	0.28	0.45	0.34
	$E_{\mathrm{c}}^{\mathrm{2g}}$	−0.12	−0.18	−0.32	−0.25

on the energy components in RPA calculations based on KS/PBE, semi/PBE and HF orbitals in paper P2. The energy components of the remaining methods were not discussed in our published papers. Understanding the convergence of the energy components in all studied methods is important. Therefore, we summarize the results of the basis set convergence of the energy components in all studied methods in Table 3.1.

Referencing Table 3.1 for KS/PBE orbitals, we see that the energy differences between AVQZ and AVTZ of the mean field energies (DFT, EXX, and RSE) remain consistently below 0.1 kJ/mol, except for the DFT energy in some cases. For example, the DFT(KS/PBE) energy difference between AVQZ and AVTZ reaches up to 0.25 kJ/mol for acetylene/II. This indicates that the 2-body DFT energies converge more slowly with the basis set compared to the other mean field energies. For the other orbital inputs studied, the convergence of the mean field energies with the basis set is similar to that observed for KS/PBE orbitals.

The components of the correlation energy (RPA, MP2, MP3, and CCSD) depend more strongly on the basis-set size, as expected. The values of these energy components obtained with AVQZ differ from those calculated with AVTZ in the range of 0.1 to 1.5 kJ/mol. The basis set convergence of the RPA correlation energies depends on DFT orbitals used. Specifically, the RPA values based on KS/SCAN and KS/SCAN0 orbitals converge more rapidly than those based on KS/PBE and KS/PBE0 orbitals (see Table 3.1). Among the correlation energies based on HF orbitals, CCSD exhibits the fastest convergence, while RPA converges the slowest. This is due to the absence of second-order exchange energy in the RPA calculations. This observation indicates that we need to use large basis set or extrapolate to CBS to obtain the RPA correlation energy converged with the basis set.

The scaling procedure proposed by Knizia et al. [261] has been shown to effectively reduce the basis-set size dependence of the (T) component [44]. We applied this procedure to the 2-body (T) calculations. The results show that the basis set convergence of (T(scaled)) component is faster than that of correlation energies discussed above. Indeed, the energy differences between AVQZ and AVTZ for this component are around only 0.05 kJ/mol or less (see Table 3.1).

The SOSEX and 2g corrections demonstrate a faster convergence than the RPA correlation energies, but slower than the mean field energies. Their energy differences between AVQZ and AVTZ fall within the range of 0.1 to 0.5 kJ/mol (see Table 3.1). These corrections converge faster for KS/PBE orbitals than for HF orbitals. Interestingly, the addition of the SOSEX corrections results in some cancellations in the basis set errors for the RPA correlation energy. This makes the RPA+SOSEX energy less sensitive to the basis set than the RPA correlation energy.

As discussed above, the energy components obtained from the 2-body calculations depend on the basis-set size, especially for the correlation energies. The question that arises now is how the convergence of the energy components with the basis set is influenced by distance. To clarify this point, we divided dimers into two groups: a proximate group ($r < 10$ Å) and a distant group ($10 < r < r_{\text{cut}}$). It is evident from Fig. 3.1 that the differences between AVQZ and AVTZ arise primarily at the proximate dimers, while these differences are close to zero for the distant dimers. This result may be understood as follows: as the distance be-

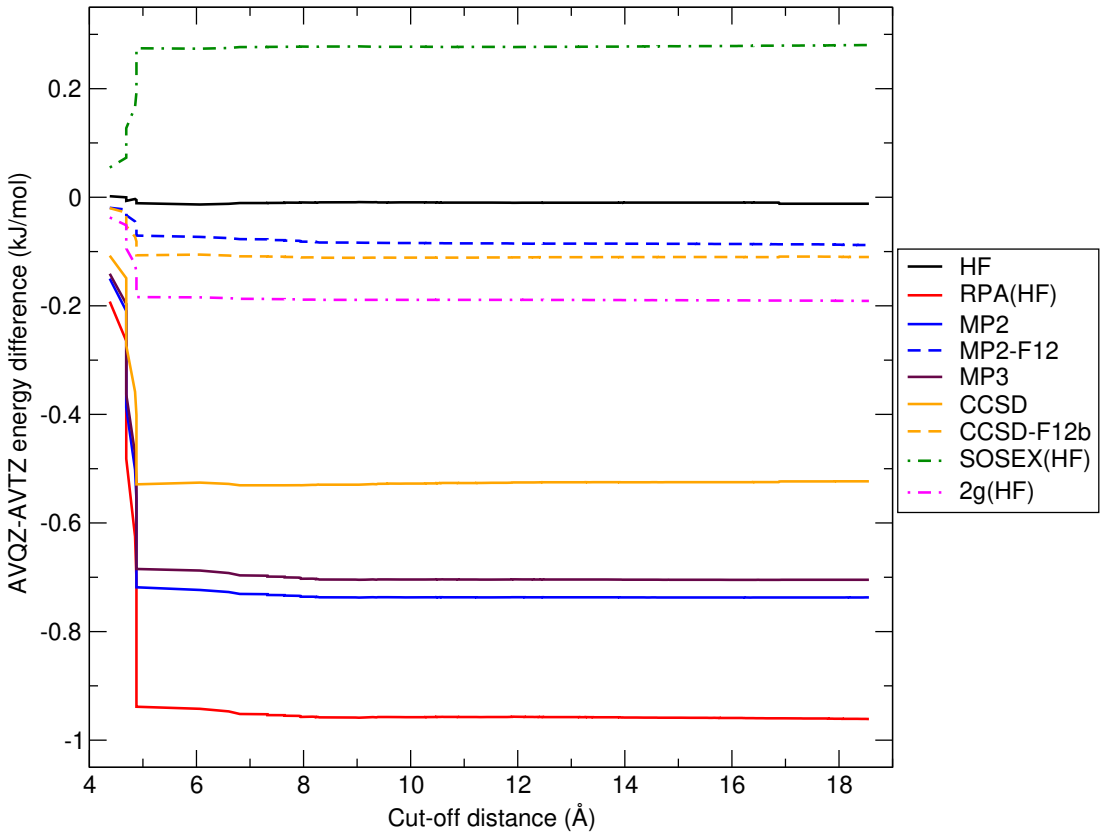

Figure 3.1: Cut-off distance dependence of the energy differences between the AVQZ and AVTZ basis sets obtained from the 2-body calculations for ethylene.

tween molecules increases, the strength of interactions decreases and perturbing potential of other molecule becomes more homogeneous so that smaller basis sets are required for the distant dimers compared to the proximate dimers. Consequently, one can use small basis sets, such as AVTZ or even AVDZ, for the distant dimers. This can lead to the significant savings of CPU time as the number of distance dimers is large.

Finally, we discuss the importance of the CABS corrections and F12 approach in improving the basis set convergence of the HF and correlation energies, respectively. As can be seen in Table 3.1, the HF+CABS values change only marginally, by less than 0.01 kJ/mol, when going from AVTZ to AVQZ basis set in comparison with the change of HF values without corrections (0.01–0.07 kJ/mol). Therefore, the use of the CABS corrections is beneficial in calculating HF component. In other words, employing CABS corrections allows us to obtain converged HF energy with a small basis set. The incorporation of the F12 corrections significantly improves the basis set convergence for the MP2 and CCSD correlation energies. Specifically, the differences between AVQZ and AVTZ for MP2-F12 and CCSD-F12b are reduced to below 0.1 kJ/mol for ethane and ethylene. However, notable differences (0.4–0.6 kJ/mol) observed for the two forms of acetylene arise from numerical errors when using the AVQZ basis set with the F12 method. This issue may come from the need to sum contributions of a large number of fragments, and was also observed before [23, 202]. Overall, the CABS and F12 corrections provide valuable means to improve the basis set convergence of the studied en-

ergy components, but numerical noise needs to be considered when using these corrections.

3.1.1.2 Non-additive terms

It was previously shown that the basis set convergence of the non-additive terms is considerably faster than that of the two-body term [23, 44, 208]. Consequently, one can use a small basis set to obtain converged values of the non-additive terms. However, the basis set convergence of the non-additive DFT energies has not yet been studied. Moreover, the convergence behavior of the non-additive RPA correlation energies in molecular solids with the basis set remains an open question. Similar to the 2-body terms, we discussed the basis set convergence of several energy components in our works. Here we will extend the discussion to include the energy components in all studied methods. The data for the basis set convergence are summarized in Table 3.2 for the 3-body term and in Table 3.3 for the 4-body term.

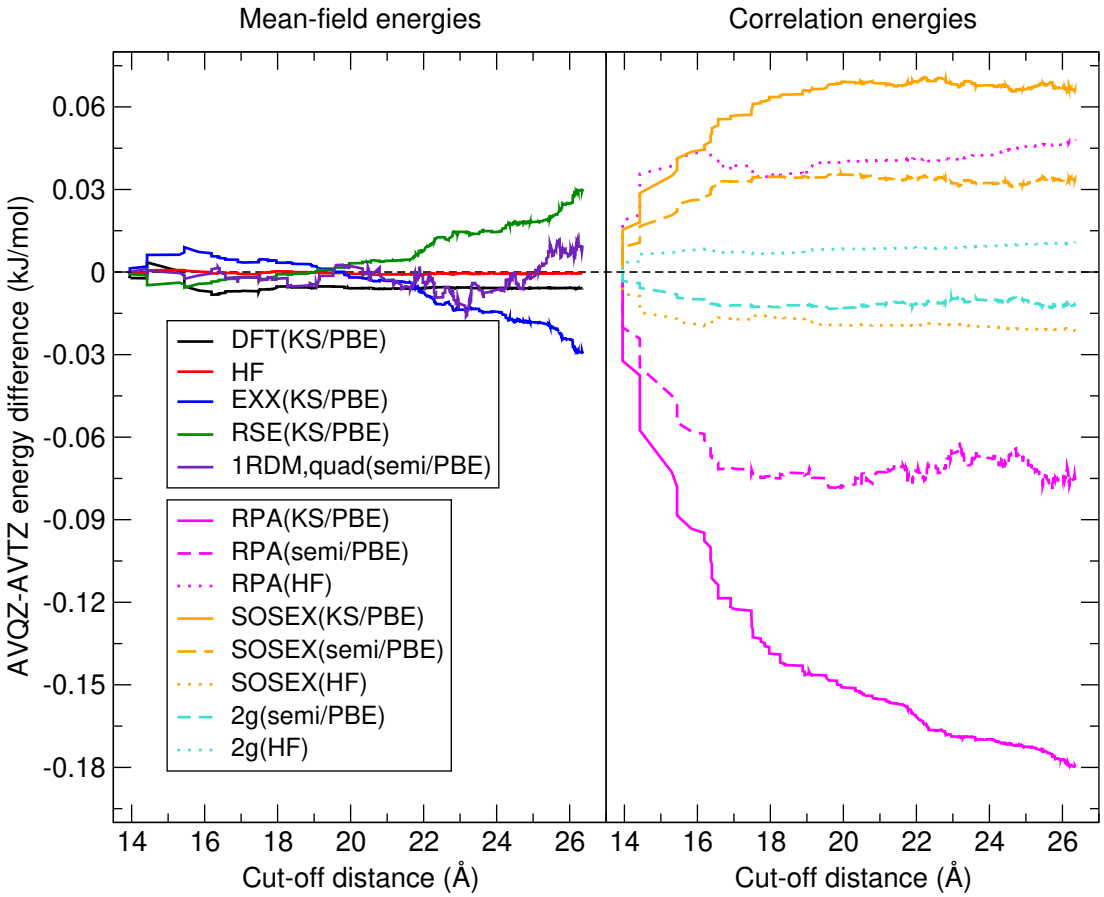

Figure 3.2: Cut-off distance dependence of the energy differences between the AVQZ and AVTZ basis sets obtained from the 3-body calculations for ethylene.

Let us first examine the basis set convergence of the mean field energies. Observing Tables 3.2 and 3.3 for KS/PBE orbitals, it is evident that the errors in the 3-body and 4-body mean field energies between the AVQZ and AVTZ basis sets are almost negligible, except for some cases. For example, large differences (around 0.1 kJ/mol) are observed in the case of the 3-body DFT and EXX energies

Table 3.2: The basis set convergence of the energy components obtained from the 3-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets. The energy components (E^{DFT}, E^{EXX}, and E^{RSE}) do not depend on the transformation from KS/PBE orbitals to semi/PBE orbitals, and are thus identical.

Inputs	Components	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	E^{DFT}	−0.03	−0.01	0.04	0.07
	E^{EXX}	0.00	−0.03	0.01	0.02
	E^{RSE}	0.00	0.03	0.00	−0.02
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.17	−0.18	−0.16	−0.14
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	0.08	0.07	0.07	0.05
semi/PBE	E^{DFT}	−0.03	−0.01	0.04	0.07
	E^{EXX}	0.00	−0.03	0.01	0.02
	E^{RSE}	0.00	0.03	0.00	−0.02
	$E^{\mathrm{1RDM,quad}}$	0.03	0.00	0.00	−0.05
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.14	−0.07	−0.06	0.00
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	0.05	0.03	0.03	0.03
	$E_{\mathrm{c}}^{\mathrm{2g}}$	−0.03	−0.01	−0.01	0.01
KS/SCAN	E^{DFT}	0.05	−0.01	0.09	0.02
	E^{EXX}	0.02	−0.03	−0.08	−0.05
	E^{RSE}	−0.02	0.03	0.08	0.06
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.34	−0.28	−0.05	−0.06
KS/PBE0	E^{DFT}	−0.03	−0.01	0.05	0.08
	E^{EXX}	0.01	0.00	0.00	−0.01
	E^{RSE}	0.00	0.00	0.01	0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.10	−0.09	−0.06	−0.07
KS/SCAN0	E^{DFT}	0.02	−0.03	0.04	0.01
	E^{EXX}	0.03	−0.01	−0.04	−0.03
	E^{RSE}	−0.03	0.01	0.04	0.04
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.21	−0.16	−0.03	−0.04
HF	E^{HF}	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	0.03	0.05	0.05	0.03
	$E_{\mathrm{c}}^{\mathrm{MP2}}$	–	0.01	–	–
	$E_{\mathrm{c}}^{\mathrm{MP2-F12}}$	–	0.01	–	–
	$E_{\mathrm{c}}^{\mathrm{MP3}}$	0.01	0.02	0.05	0.03
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	−0.01	−0.02	−0.03	−0.01
	$E_{\mathrm{c}}^{\mathrm{2g}}$	0.00	0.01	0.01	0.00

Table 3.3: The basis set convergence of the energy components obtained from the 4-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets.

Inputs	Components	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	E^{DFT}	0.00	0.01	−0.01	−0.01
	E^{EXX}	0.01	−0.01	0.00	0.00
	E^{RSE}	0.00	0.00	0.01	−0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	0.07	0.08	0.11	0.08
KS/SCAN	E^{DFT}	−0.07	0.00	0.01	0.02
	E^{EXX}	0.00	0.00	0.03	0.00
	E^{RSE}	0.00	0.00	−0.02	−0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	0.14	0.10	0.00	0.00
KS/PBE0	E^{DFT}	0.00	0.01	0.00	−0.01
	E^{EXX}	0.00	0.00	0.00	0.00
	E^{RSE}	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	0.04	0.05	0.07	0.04
KS/SCAN0	E^{DFT}	−0.05	0.00	0.02	0.01
	E^{EXX}	−0.01	0.00	0.02	0.01
	E^{RSE}	0.00	0.00	−0.02	−0.01
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	0.08	0.08	0.01	0.00
HF	E^{HF}	–	0.00	–	–
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	–	0.00	–	–

based on KS/SCAN for acetylene/I (see Table 3.2). However, these significant errors can stem from numerical noise. This can be seen in Fig. 3.2 as an example, where the difference in the 3-body DFT (KS/PBE) energy is nearly zero at short distances but begins to increase at larger distances. The behavior of the basis set convergence remains consistent for the non-additive mean field energies based on other inputs. Overall, the non-additive mean field energies can achieve converged values with small basis sets, such as AVTZ or even AVDZ, without the need for extrapolations.

We now turn to discuss the correlation energies. The errors of the 3-body MP2 and MP3 correlation energies between AVQZ and AVTZ are very small (below 0.02 kJ/mol), as shown in Table 3.2. However, the data show that the non-additive RPA correlation energies exhibit relatively slow convergence with the basis set size. For example, the 3-body RPA(KS/PBE) correlation energies obtained with AVQZ differ from those obtained with AVTZ by around 0.1–0.2 kJ/mol (Table 3.2). These errors are significant for the non-additive contributions, and it is necessary to extrapolate the non-additive RPA correlation energies to CBS to achieve converged values.

The basis set convergence of the non-additive RPA correlation energies is influenced by the choice of orbitals. Indeed, the 3-body RPA correlation energy based on HF orbitals shows a faster convergence compared to that based on DFT orbitals, see Fig. 3.2. Moreover, the convergence can be also affected by characteristics of the considered systems. It is observed that the convergence is quicker for the two forms of acetylene than for ethane and ethylene. For

instance, the error in the 3-body RPA correlation energy based on KS/SCAN is only 0.05 kJ/mol for acetylene/I, but reaches approximately 0.4 kJ/mol for ethane (Table 3.2). This smaller error for acetylene is observed for each trimer contributing to the total 3-body energies. Thus, the non-additive RPA correlation energies tend to converge more rapidly with the basis set size for the studied systems with π electrons.

Finally, we examine the basis set convergence of the 3-body RPA corrections (SOSEX and 2g). These energy components show very little dependence on the basis-set size, changing by less than 0.10 kJ/mol between the AVTQZ and AVTZ basis sets (Table 3.2). Similar to the RPA correlation energy, these corrections converge faster when based on HF orbitals compared to KS/PBE and semi/PBE orbitals. Therefore, the use of HF orbitals can be beneficial in enhancing basis set convergence of the RPA correlation energy and its corrections. Moreover, the slow basis set convergence of the 3-body RPA correlation energies can be reduced partly when adding the SOSEX corrections to the RPA correlation energies. This is again due to the opposite basis set convergence of the second-order direct and exchange terms [268, 269].

3.1.1.3 Used basis sets and uncertainties

Based on the convergence behavior of the energy components with the basis set discussed, we decided to use the basis set settings for the energy components shown in Table 3.4. We attempted to use the basis sets that are sufficiently large to achieve converged values of the energy components. However, due to computational cost, we were unable to reach converged values for the energy components exhibiting slow basis set convergence to within 0.01 kJ/mol. Therefore, we now try to estimate the errors in the energy components associated with the basis-set convergence problem.

Table 3.4: The basis sets used for obtaining converged values of the energy components for all the considered systems.

Energy components	2-body	3-body	4-body
E^{HF}	AVQZ	AVTZ	AVTZ
E^{DFT}	AVQZ	AVQZ	AVQZ
E^{EXX}	AVQZ	AVQZ	AVQZ
E^{RSE}	AVQZ	AVQZ	AVQZ
$E^{\mathrm{1RDM,quad}}$	AVQZ	AVQZ	-
$E_{\mathrm{c}}^{\mathrm{RPA}}$	AVTZ→AVQZ	AVTZ→AVQZ	AVTZ→AVQZ
$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	AVTZ→AVQZ	AVTZ→AVQZ	-
$E_{\mathrm{c}}^{\mathrm{2g}}$	AVTZ→AVQZ	AVTZ→AVQZ	-
$E_{\mathrm{c}}^{\mathrm{MP2}}$	AVTZ→AVQZ	AVTZ	AVTZ
$E_{\mathrm{c}}^{\mathrm{MP2-F12}}$	AVTZ→AVQZ	AVTZ	AVTZ
$E_{\mathrm{c}}^{\mathrm{MP3}}$	AVTZ→AVQZ	AVTZ→AVQZ	-
$E_{\mathrm{c}}^{\mathrm{CCSD}}$	AVTZ→AVQZ	AVTZ	AVDZ
$E_{\mathrm{c}}^{\mathrm{CCSD-F12b}}$	AVTZ→AVQZ	AVTZ	AVDZ
$E_{\mathrm{c}}^{\mathrm{T}}$	AVQZ	AVTZ	AVDZ

The errors of the mean field energies for all n-body contributions can be safely

neglected when AVQZ or AVTZ basis set was used. The errors of the 3-body and 4-body correlation energies can also be considered negligible, except for the RPA correlation energies based on DFT orbitals. Based on their convergence behavior, we can estimate their uncertainties to be less than 0.05 kJ/mol.

The main errors arising from the basis set convergence problem are observed for the 2-body correlation energies. As discussed previously, the basis set convergence of the 2-body correlation energies depends on examined methods. However, performing the calculations with all methods using large basis sets is computationally demanding. Therefore, for a general estimation, we only performed the MP2 calculations for ethylene to obtain values with various basis sets presented in Table 3.5.

Table 3.5: Basis set convergence of the 2-body MP2 correlation energies in kJ/mol for ethylene. Here the energies are given to three decimal digits to be able to show small changes between them.

Basis sets	MP2	MP2-F12
AVDZ	−31.070	−34.191
AVTZ	−33.764	−34.888
AVQZ	−34.051	−35.022
AV5Z	−34.753	−34.994
AVDZ→AVTZ	−34.899	−34.995
AVTZ→AVQZ	−35.039	−35.004
AVQZ→AV5Z	−35.017	−35.003

We note that all 2-body correlation energies of the studied methods were obtained with AVTZ→AVQZ extrapolation (see Table 3.4), except for (T) component, where the AVQZ basis set was used. One can see from Table 3.5 that the difference in the 2-body MP2 correlation energy between AVTZ/AVQZ and AVQZ/AV5Z is very small (below 0.03 kJ/mol). Therefore, it is reasonable to assume that uncertainties arising from the basis set convergence issue are negligible for our purposes, around a few hundredths of kJ/mol. The exception here are the 2-body RPA correlation energies, which converge more slowly with the basis set size compared to the correlation energies based on wavefunction methods. Based on their convergence behavior, we estimate their errors to be around 0.05 kJ/mol or less.

3.1.2 Cut-off distance

The values of n-body contributions depend on the number of fragments included in each contribution (dimer, trimer, ...), which is theoretically infinite. However, in practice, only a limited number of fragments (within some cut-off distance) are typically considered. In other words, we truncate the calculations at distances that are sufficient to obtain converged values of the energy components. This is because interactions at large distances have minimal contribution on the overall energy [23], and the calculations involving such distances can introduce numerical noise [23, 202]. The cut-off distances and the number of symmetry inequivalent fragments for each system used in our calculations are listed in Ta-

ble 3.6. Understanding how the energy components converge with respect to the cut-off distance is crucial for determining appropriate cut-off values for calculating energy components accurately. We will discuss this based on paper P1 in this section, starting with the 2-body terms, and then with the non-additive terms.

Table 3.6: Cut-off distance (r_{cut}, in Å) and corresponding number (N) of symmetry inequivalent dimers, trimers, and tetramers within the selected cut-off distance for the MBE calculations.

Systems	2-body		3-body		4-body	
	r_{cut}	N	r_{cut}	N	r_{cut}	N
Ethane	19.5	436	25.2	991	34.6	200
Ethylene	18.6	428	26.3	1672	33.1	202
Acetylene/I	24.4	1174	27.0	2875	31.8	282
Acetylene/II	24.8	1094	27.4	2655	32.1	164

In this section, we divide the energy components into two groups. The first group comprises the mean field energies (DFT, EXX, HF, RSE, and 1RDM,quad) and SOSEX corrections. The second group includes the correlation energies (RPA, MP2, MP3, CCSD, and T) and 2g corrections. This classification is based on their similar convergence behavior with respect to the cut-off distance. The energy components in the first group are expected to converge quickly with the cut-off distance, while those in the second group converge slowly due to long-range correlation interactions. We also note that the convergence behavior of each energy component in the RPA calculations with the cut-off distance remains consistent for all the examined orbitals.

3.1.2.1 Two-body terms

As previously noted, some energy components exhibit similar convergence behavior with distance. Hence, we only present the convergence of selected energy components as representative examples. Specifically, we show the convergence of DFT, EXX, and RSE for the first group, and RPA for the second group.

We start to discuss the convergence of the energy components in the first group. One can see from Fig. 3.3(a) that the convergence of the DFT energy is remarkably rapid for ethane. However, an oscillatory convergence trend is observed for the two forms of acetylene. This behavior can be attributed to differences in electrostatic moments of molecules, especially the quadrupole moment. Specifically, ethane has a zero moment, whereas for acetylene, it is approximately 4 a.u [270, 271]. Similar oscillatory convergence patterns were observed for other systems as well [23]. The convergence behavior with the cut-off distance of other energies in the first group is similar to that of the DFT energy. However, it is remarkable that the convergence of the singles corrections (RSE and 1RDM,quad) exhibits relatively little dependence on the systems compared to the other mean field energies (see Fig 3.3(c)). These energy components show a very rapid convergence with the cut-off distance, and change only 0.01 kJ/mol for ethane and ethylene, and 0.03 kJ/mol for the two forms of acetylene between 8 Å and cut-off used.

Table 3.7: The values of the energy components (kJ/mol) obtained from the 2-body calculations for distant dimers ($10 < r < r_{cut}$) in different approaches.

Inputs	Components	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	E^{DFT}	0.00	0.05	−0.02	0.10
	E^{EXX}	0.00	0.06	−0.07	0.15
	E^{RSE}	0.00	0.00	0.05	−0.02
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.53	−0.51	−0.46	−0.37
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	−0.04	−0.03	−0.03	0.01
Semi/PBE	E^{DFT}	0.00	0.05	−0.02	0.10
	E^{EXX}	0.00	0.06	−0.07	0.15
	E^{RSE}	0.00	0.00	0.05	−0.02
	$E^{\mathrm{1RDM,quad}}$	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.41	−0.39	−0.31	−0.30
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	−0.02	−0.01	−0.01	0.01
	$E_{\mathrm{c}}^{\mathrm{2g}}$	−0.10	−0.10	−0.07	−0.08
KS/SCAN	E^{DFT}	0.00	0.05	−0.03	0.11
	E^{EXX}	0.00	0.06	−0.03	0.13
	E^{RSE}	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.51	−0.46	−0.40	−0.33
KS/PBE0	E^{DFT}	0.00	0.05	−0.03	0.11
	E^{EXX}	0.00	0.06	−0.03	0.12
	E^{RSE}	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.48	−0.45	−0.39	−0.33
KS/SCAN0	E^{DFT}	0.00	0.05	−0.03	0.11
	E^{EXX}	0.00	0.06	−0.03	0.11
	E^{RSE}	0.00	0.00	0.00	0.00
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.46	−0.43	−0.38	−0.34
HF	E^{HF}	0.00	0.06	−0.03	0.13
	$E^{\mathrm{HF+CABS}}$	0.00	0.06	−0.03	0.13
	$E_{\mathrm{c}}^{\mathrm{RPA}}$	−0.37	−0.37	−0.35	−0.29
	$E_{\mathrm{c}}^{\mathrm{MP2}}$	−0.61	−0.63	−0.62	−0.53
	$E_{\mathrm{c}}^{\mathrm{MP2-F12}}$	−0.61	−0.62	−0.60	−0.55
	$E_{\mathrm{c}}^{\mathrm{MP3}}$	−0.57	−0.55	−0.45	−0.38
	$E_{\mathrm{c}}^{\mathrm{CCSD}}$	−0.51	−0.47	−0.40	−0.35
	$E_{\mathrm{c}}^{\mathrm{CCSD-F12b}}$	−0.51	−0.47	−0.40	−0.35
	$E_{\mathrm{c}}^{\mathrm{T}}$	−0.10	−0.10	−0.09	−0.08
	$E_{\mathrm{c}}^{\mathrm{SOSEX}}$	−0.02	−0.01	0.01	0.01
	$E_{\mathrm{c}}^{\mathrm{2g}}$	−0.09	−0.09	−0.08	−0.07

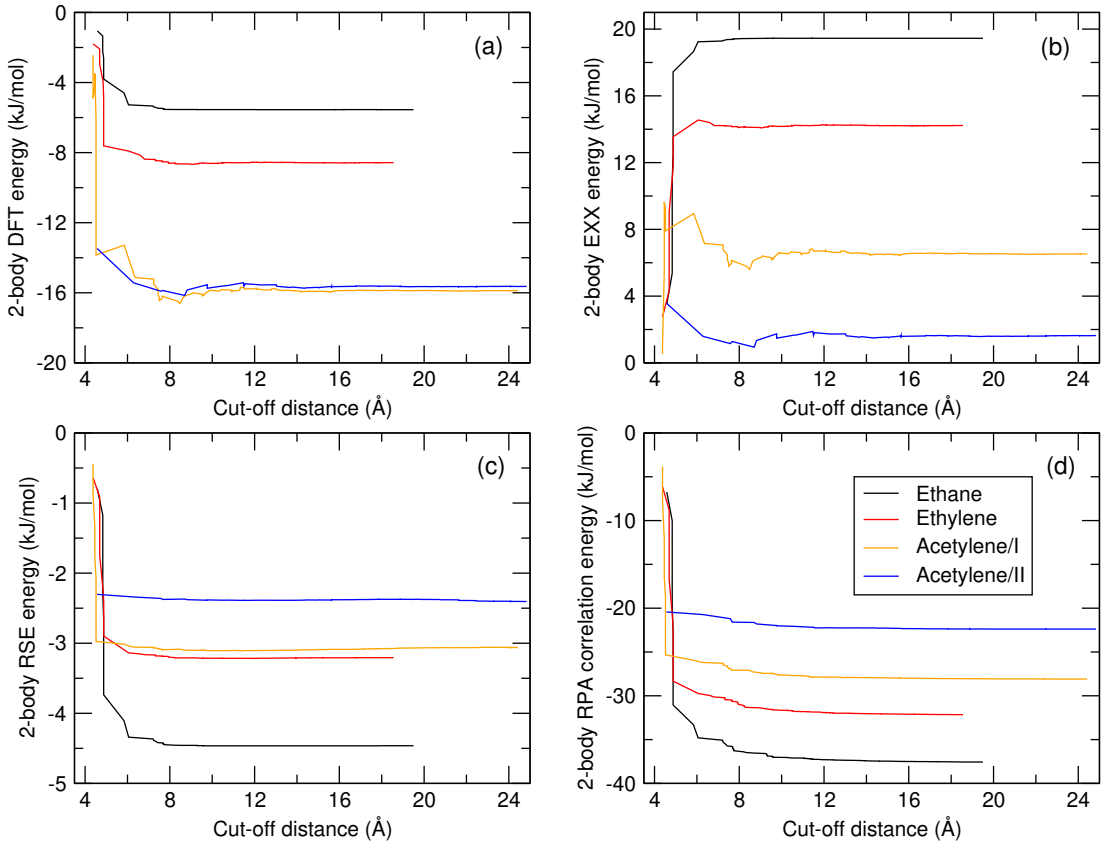

Figure 3.3: Cut-off distance convergence of the 2-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used.

We now shift our focus to the convergence of the energy components in the second group: correlation energies and 2g corrections. As depicted in Fig. 3.3(d), the convergence trend of the RPA correlation energy is similar for all the considered systems. Specifically, contributions at small cut-off distances dominate, and minimal or no oscillations are observed for larger cut-off distances. The same convergence behavior is observed for the remaining energy components in the second group due to the need to describe the long-range correlation interactions.

Let us examine in a more detail the contributions of the energy components at large distances presented in Table 3.7. The data show that the contributions of the singles corrections (RSE and 1RDM,quad) and SOSEX components at large distances are close to zero. This implies that these energy components converge fast with distance and their converged values with the cut-off distance can be obtained at short cut-off values. The HF and EXX energies, although small, are non-negligible for the systems with π interactions. For instance, the contribution of the EXX(KS/PBE) component is up to 0.15 kJ/mol for acetylene/II (see Table 3.7). This is mainly due to differences in electrostatic moments of the considered molecules as discussed above. For the energy components in the second group (correlation energies and 2g corrections), we observe a slow convergence with the cut-off distance, which is expected because the correlation interactions decay proportionally to $-r^{-6}$ with the intermolecular distance r. The contributions of the correlation energies at large distance are between -0.30 kJ/mol and -0.60 kJ/mol. The contributions of the 2g energies are smaller, near 0.1 kJ/mol,

for all the considered systems. Therefore, obtaining converged values of the correlation energy components and 2g corrections with the cut-off distance requires performing calculations at very large distances or extrapolating with distance.

3.1.2.2 Non-additive terms

We now move to discuss the energy components obtained from the 3-body and 4-body calculations. Similar to the 2-body terms, we will focus on the cut-off distance convergence of the DFT, EXX, RSE, and RPA components as representative examples. These are shown in Fig. 3.4 for the 3-body terms and in Fig. 3.5 for the 4-body terms.

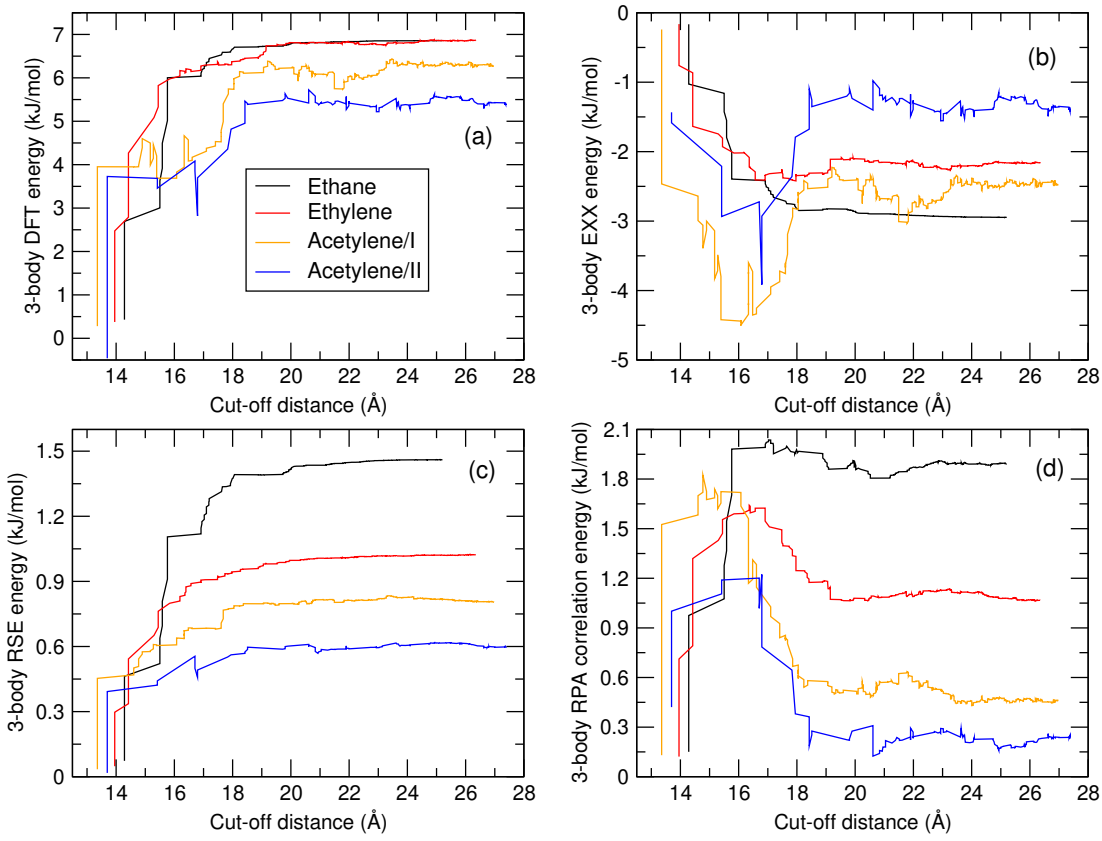

Figure 3.4: Cut-off distance convergence of the 3-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used.

Let us start with the 3-body terms. One can see from Fig. 3.4(b) that the convergence of the 3-body EXX energy is rather fast for ethane. However, notable fluctuations with both negative and positive terms occur for the two forms of acetylene with distances below 20 Å that lead to changes of several kJ/mol. Beyond that distance, the 3-body EXX energy tends to stabilize and converge within a few tenths of kJ/mol. This means that the convergence of the 3-body EXX energy with the cut-off distance clearly depends on the magnitude of the electrostatic moments of the considered systems. Similar behavior is observed for the other energies in the first group (DFT, HF, RSE, 1RDM,quad, and SOSEX corrections). The convergence of the 3-body singles corrections (RSE and 1RDM,quad) is observed to have minimal dependence on the systems when compared to other mean field energies, similar to that observed for the 2-body terms.

The contributions of these corrections beyond 20 Å are less than 0.05 kJ/mol for all the considered systems.

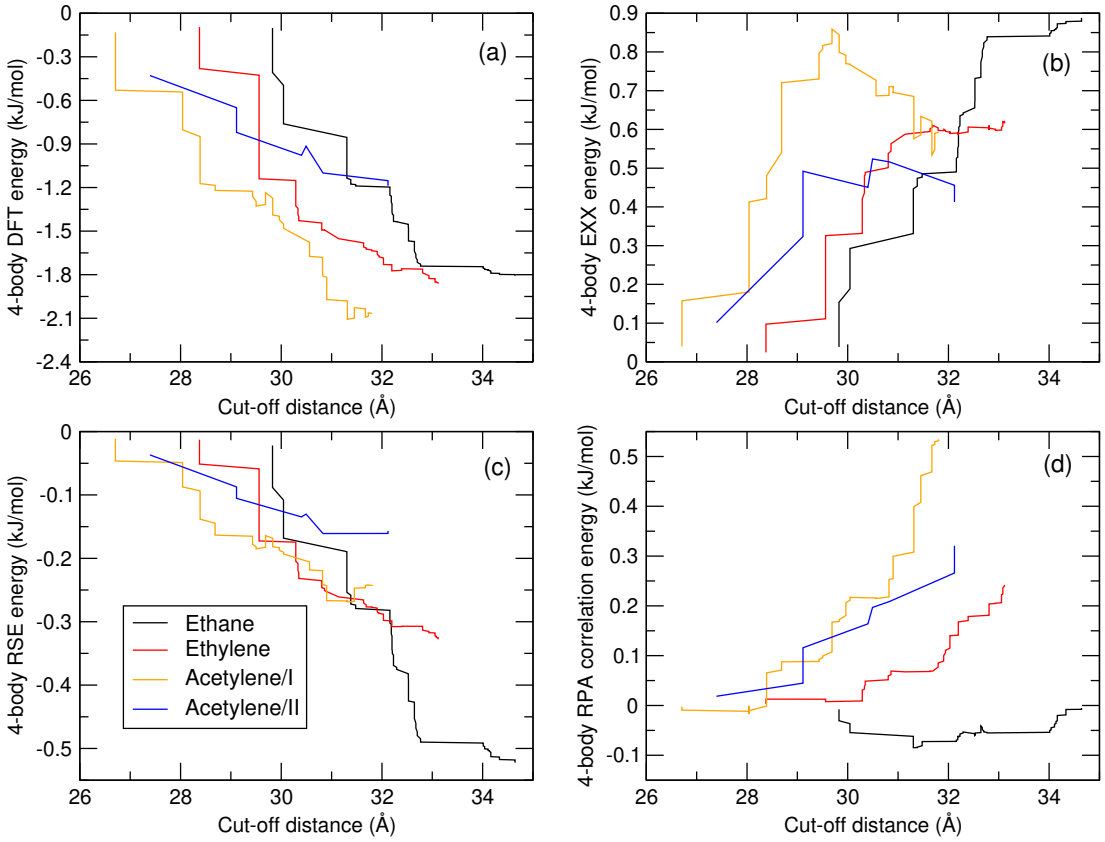

Figure 3.5: Cut-off distance convergence of the 4-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used.

One can see from Fig. 3.4(d) that the convergence of the 3-body RPA correlation energies with the cut-off distance is almost similar for all the considered systems. The dominant contributions occur within distances below 20 Å, and beyond that threshold, they change by less than 0.2 kJ/mol. For the other energies within the second group (MP2, CCSD, (T) and 2g corrections), the same convergence behavior as in the case of the RPA correlation energy is observed.

We now turn to consider the 4-body terms. From Fig. 3.5, it can be observed that all 4-body energy components are not completely converged with the cut-off distance used. The convergence is complicated, and significant changes are observed for all the energy components. We have also tested that extending the cut-off by several Å can lead to numerical noise and still changes the values of energy components by few tenths of kJ/mol. This shows that it is difficult to obtain converged values of the 4-body energy components with the cut-off distance. Further investigation is needed to gain a better understanding of this issue.

3.1.2.3 Uncertainties

The use of the cut-off distance typically introduces errors in the energy components, and we now try to estimate these errors. This is important for gaining a

better understanding when comparing MBE results with PBC in the next chapter. Here we only estimate the errors in the energy components for the 2-body and 3-body calculations because the 4-body energy components are not completely converged with the cut-off distance, as discussed above.

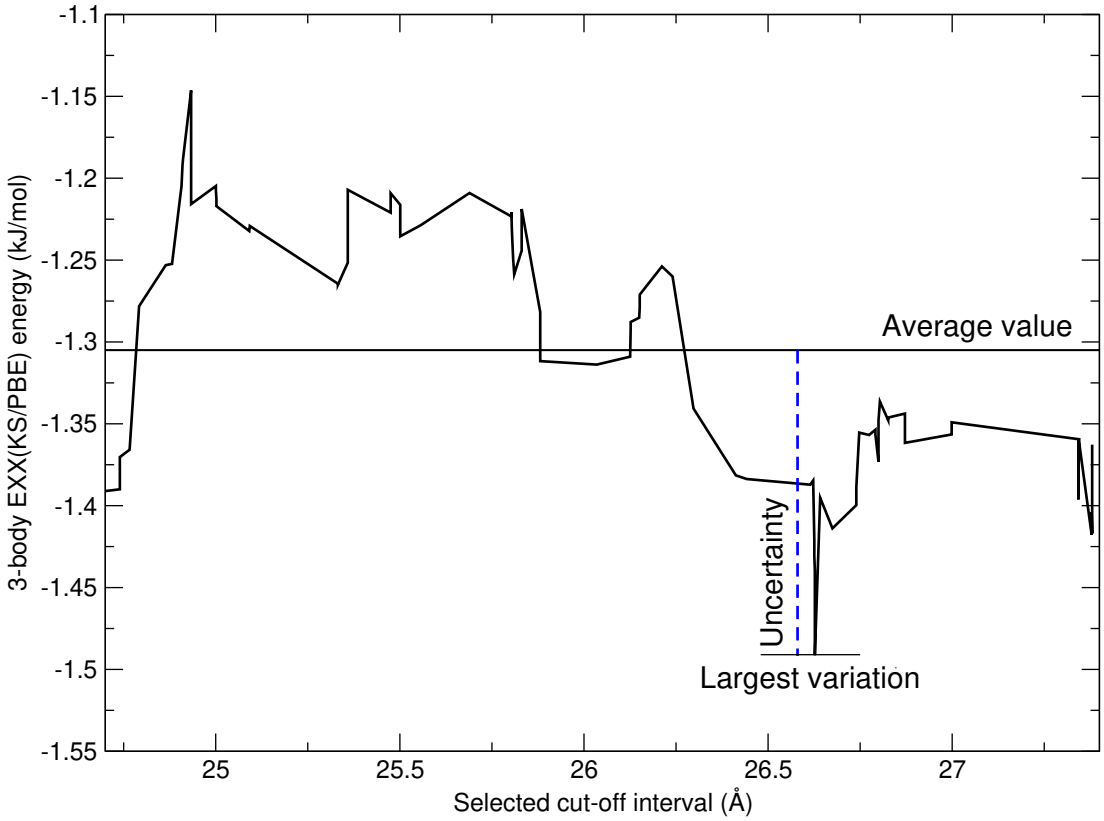

Figure 3.6: Method for estimating the uncertainty of the 3-body EXX contribution based on KS/PBE of acetylene/II.

To estimate the uncertainty of an energy component, we first calculate the average value of this energy component within a specified interval, ranging from a fixed cut-off to the largest cut-off. The selected intervals for all systems are summarized in Table 3.8. Then, we determine the uncertainty by examining the variations in energy within the selected interval relative to its average value. This procedure is illustrated more clearly in Fig. 3.6, and is similar to that used in Ref. [23].

The uncertainties for the energy components in the first group are provided in Table 3.9, while those for the energy components in the second group are listed in Table 3.10. We note that the errors in the energy components in the RPA calculations displayed in Tables 3.9 and 3.10 are based on KS/PBE orbitals. However, these errors remain nearly constant, with changes of less than 0.02 kJ/mol, when other orbitals are employed.

For the energy components in the first group, the uncertainties of the 2-body terms are safely negligible (below 0.03 kJ/mol). The uncertainties are larger for the 3-body terms. However, these larger errors are only observed for the DFT, EXX, and HF energies of the systems with delocalized electrons. For instance, an uncertainty of 0.19 kJ/mol is noted for the 3-body EXX energy of acetylene/II

Table 3.8: Cut-off distance interval (R, in Å) and corresponding number (N) of dimers and trimers selected to estimate the uncertainties of energy components.

Systems	2-body		3-body	
	R	N	R	N
Ethane	15.5−19.5	210	22.2−25.2	598
Ethylene	14.6−18.6	224	23.3−26.3	856
Acetylene/I	20.4−24.4	516	24.0−27.0	1630
Acetylene/II	20.8−24.8	444	24.7−27.4	1473

Table 3.9: The estimated uncertainties (in kJ/mol) of the energy components in the first group of the methods for all the considered systems.

Components	Contributions	Ethane	Ethylene	Acetylene/I	Acetylene/II
DFT	2-body	0.00	0.01	0.02	0.01
	3-body	0.02	0.07	0.06	0.18
EXX	2-body	0.00	0.01	0.02	0.03
	3-body	0.02	0.06	0.10	0.19
HF	2-body	0.00	0.01	0.02	0.02
	3-body	0.00	0.07	0.10	0.20
RSE	2-body	0.00	0.00	0.00	0.02
	3-body	0.01	0.00	0.01	0.01
1RDM,quad	2-body	0.00	0.00	0.01	0.00
	3-body	0.03	0.00	0.01	0.01
SOSEX	2-body	0.00	0.01	0.00	0.01
	3-body	0.02	0.01	0.02	0.04

Table 3.10: The estimated uncertainties (in kJ/mol) of the energy components in the second group of the methods for all the considered systems.

Components	Contributions	Ethane	Ethylene	Acetylene/I	Acetylene/II
RPA	2-body	0.05	0.06	0.01	0.01
	3-body	0.02	0.04	0.03	0.05
MP2-F12	2-body	0.05	0.07	0.02	0.02
	3-body	0.01	0.01	0.03	0.04
MP3	2-body	0.05	0.07	0.01	0.01
	3-body	0.08	0.02	0.07	0.06
CCSD-F12b	2-body	0.04	0.06	0.01	0.01
	3-body	0.05	0.01	0.05	0.05
T	2-body	0.01	0.01	0.00	0.00
	3-body	0.00	0.01	0.00	0.00
2g	2-body	0.01	0.01	0.00	0.00
	3-body	0.01	0.01	0.01	0.03

(see Table 3.9). The uncertainties of the remaining energies in the first group for the 3-body terms can generally be considered insignificant.

For the energy components in the second group, the uncertainties of (T) and 2g energy components for both 2-body and 3-body terms are mostly below 0.01 kJ/mol. The exception is the 3-body 2g energy of acetylene/II, which has an uncertainty of 0.03 kJ/mol. The uncertainties of the other correlation energies are slightly larger but less than 0.1 kJ/mol.

3.1.3 Summary

The convergence of the energy components in examined methods obtained from the MBE approach with the basis set and the cut-off distance was discussed in detail. The key observations from these discussions can be summarized as follows:

The basis set convergence rate is different for the energy components. Extrapolations to CBS are necessary to achieve converged correlation energies, whereas sufficiently large basis sets, such as AVTZ, are adequate for obtaining converged mean field energies. The convergence of the RPA correlation energies with the basis set is comparatively slower than that of the correlation energies obtained by wavefunction methods. The choice of the examined orbitals can also influence the basis set convergence of the RPA correlation energies, with those based on HF orbitals converging faster than those based on DFT orbitals. Adding the SOSEX corrections to the RPA correlation energies can mitigate the slow basis set convergence of the RPA correlation energies because the SOSEX corrections remove the spurious second-order self-interaction terms at short distances in the RPA correlation energies. Furthermore, the basis set convergence of the energy components depends on distance, and large basis sets are mainly required for interactions at short distances, while small basis sets can be used for interactions at large distances to save computational cost.

The mean field energies and SOSEX corrections exhibit a more rapid convergence with the cut-off distance compared to the correlation energies and 2g corrections. As a result, the converged values of the mean field energies and SOSEX corrections can be obtained at short distances, while large distances or extrapolations to infinite distance are required for obtaining converged correlation energies and 2g corrections due to the need to account for long-range correlation interactions. The cut-off distance convergence of the energy components is affected by the characteristics of the considered systems. Obtaining convergence for the energy components involving the non-additive terms with the cut-off distance, particularly for the 4-body terms, is challenging and can not be completely achieved. This issue can be possibly mitigated by employing a force field to account for the contributions at large distances [235].

3.2 Periodic boundary conditions

Binding energy of molecular solids can be obtained by the PBC approach, which is generally simpler to set up compared to MBE. However, achieving a converged binding energy of a molecular solid in PBC requires extrapolating the

energy with respect to several numerical parameters, leading to a substantial increase in computational cost, and complexity of the setup. This section will focus on understanding how the energy components in the RPA and MP2 calculations converge with numerical parameters. We will start with k-points and cell volume used for the calculations of solids and molecules respectively, then with plane wave basis set, and finally with PAW potentials.

3.2.1 k-points and cell volume

The convergence of the energy components with k-points and cell volume are discussed in detail in our forthcoming work. Here we just mention the main features. For our calculations, we used a $(N_k)^3$ k-point grid for solids and a simulation box $a \times (a+1) \times (a+2)$ Å^3 for isolated molecules. The values of N_k and a were varied to assess how they affect the convergence and to find reliable settings. These tests are discussed below.

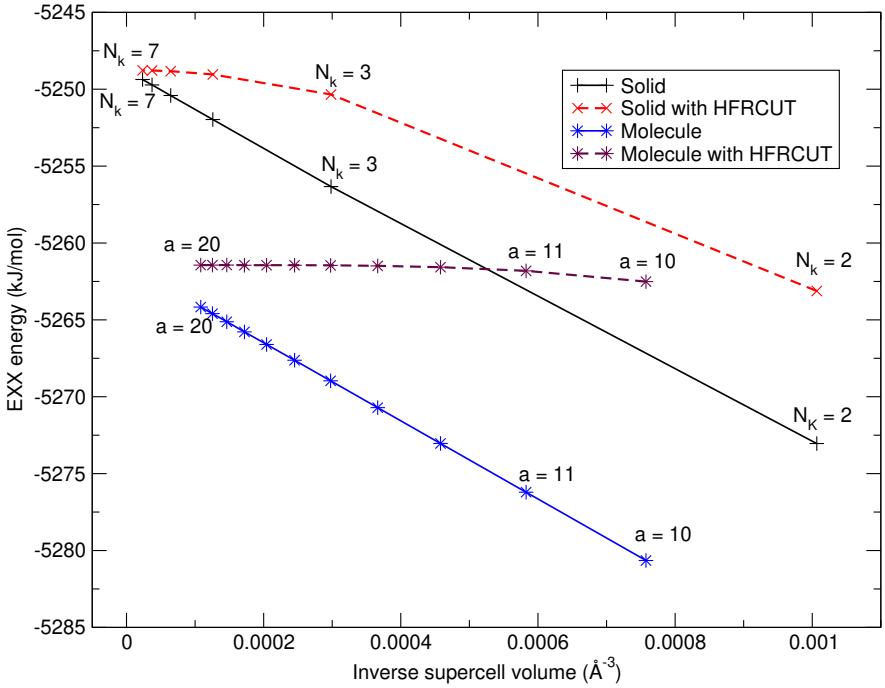

Figure 3.7: Convergence of the EXX(KS/PBE) energy for ethylene solid and isolated molecule as a function of the supercell volume without and with the use of the real-space Coulomb cut-off. For the solid, the supercell volume is obtained as the unit cell volume times the number of k-points and data between 2^3 and 7^3 k-points are shown. For the molecule, the supercell is the simulation cell. A plane-wave basis-set cut-off of 1100 eV was used.

We begin by evaluating the convergence of the mean field energies, including DFT, EXX, HF, and RSE energies, with k-points for solid calculations and cell volume for molecule calculations. The convergence of these energies with k-points and cell volume can be slow due to the presence of the Coulomb singularity in the exchange potential [272, 273], except for the DFT energy based on semilocal orbitals (KS/PBE and KS/SCAN). Therefore, it is necessary to extrapolate the energy to infinite k-points and cell volume assuming a dependence of $1/N_k^3$ for

solid and $1/V$ for molecule to obtain converged values [21]. Alternatively, one can use Coulomb cut-off technique [274] with sufficiently large number of k-points and large cell volume to reach converged energy values [23]. The advantage of this technique in improving significantly the convergence of the mean field energies with k-points and cell volume is illustrated in Fig. 3.7 as an example. Indeed, the EXX(KS/PBE) values obtained from $N_k = 3$ and $N_k = 7$ differ by 6.9 kJ/mol per molecule without the use of the technique, while this difference is reduced to only 1.5 kJ/mol when the technique is used.

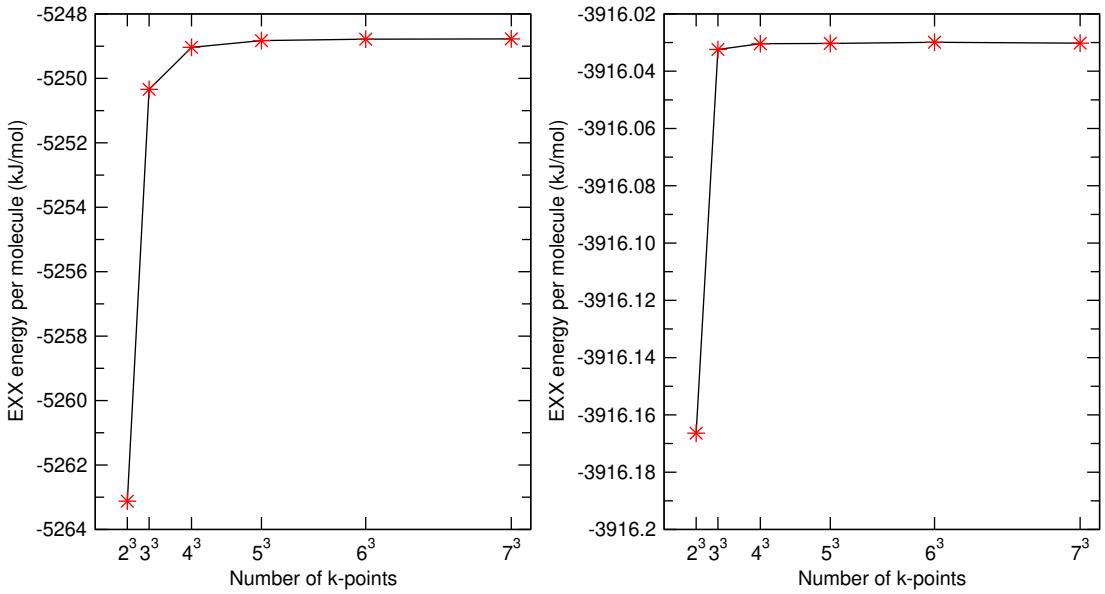

Figure 3.8: The convergence of the EXX energy for solid ethylene (left) and solid acetylene/II (right) with respect to the number of k-points. The calculations used a plane-wave basis-set cut-off of 1100 eV and the Coulomb cut-off method. Note the different scales on the y-axis. The EXX energy gives the values printed by VASP.

The convergence of the mean field energies with k-points is not similar for all the considered solids. Specifically, the convergence is notably faster for the two forms of acetylene than for ethane and ethylene, see Fig. 3.8 as an example. This difference may be attributed to variations in cell volume and symmetry among the considered systems. The exception here is the DFT energy based on KS/PBE and KS/SCAN states, where a k-points with $N_k = 4$ is sufficient for all the considered solids. For all isolated molecules considered, the convergence of the mean field energies with cell volume is almost the same.

In our calculations of the mean field energies, we decided to use the Coulomb cut-off technique to obtain converged values because this approach is reasonably simple to perform and avoids the unnecessary uncertainty that extrapolations can introduce. However, using this technique requires a sufficiently large number of k-points and a large cell volume. Here, we tried to use large number of k-points and large cell volume so that the uncertainties related to the convergence problems of k-points and cell volume for all the mean field energies can be safely negligible. These converged settings for k-points and cell volume for all the considered systems are summarized in Table 3.11. However, for the RSE calculations of solid ethane and ethylene, we applied the extrapolations. This is

because the calculations of the RSE energy for these two solids with the large number of k-points ($N_k = 6$) are computationally demanding. However, a spot check for KS/PBE and ethylene shows that the extrapolated RSE value (from $N_k = 3$ to $N_k = 4$) are within 0.01 kJ/mol of energy obtained with the cut-off scheme and $N_k = 6$.

Table 3.11: The k-point set and the cell side a in Å used for the PBC calculations. DFT, RPA are settings for KS/PBE and KS/SCAN calculations, DFTh, RPAh are settings for KS/PBE0 and KS/SCAN0 calculations, $\rightarrow$ denotes extrapolation of the energies with steps of 1 k-point grid for solids or 1 Å for isolated molecules.

Methods	Parameters	Ethane	Ethylene	Acetylene/I	Acetylene/II
DFT	N_k	4	4	4	4
	a	17	17	17	17
DFTh	N_k	6	6	4	4
	a	20	20	20	20
EXX/HF	N_k	6	6	3	3
	a	17	17	17	17
RSE	N_k	3→4	3→4	3	3
	a	10	10	10	10
RPA	N_k	4	4	4	4
	a	10	10	10	10
RPAh	N_k	2→4	2→4	2→4	2→4
	a	8→10	8→10	8→10	8→10
xMP2	N_k	3	3	2	2
	a	11	7→11	7→11	7→11
dMP2	N_k	4	4	4	4
	a	11	7→11	7→11	7→11

We now shift our discussion to the RPA correlation energy. To obtain the RPA converged values with k-points and cell volume, the extrapolations are required in case that the RPA correlation energies converge slowly with these parameters. When based on semilocal functionals, the RPA correlation energies are extrapolated to infinite number of k-points and infinite cell volume using a dependence of $1/N_k^6$ for solids and a dependence of $1/V^2$ for molecules. This convergence behavior was observed in previous work [275]. However, for the RPA correlation energies based on hybrids, the extrapolations are done using different convergence behavior: $1/N_k^3$ for solids and $1/V$ for molecules. This arises from the treatment of singularity of Hartree-Fock potential at the Γ point [272, 273]. One can see from Fig. 3.9 that the convergence of the RPA correlation energy is more rapid for KS/PBE orbitals than for KS/PBE0 orbitals. Specifically, the difference in the RPA correlation energy per molecule between $N_k = 3$ and $N_k = 4$ is very small (below 0.05 kJ/mol) when using KS/PBE orbitals, but notably larger, exceeding 0.1 kJ/mol, with KS/PBE0 orbitals. Similarly, for isolated molecules, the discrepancy between boxes with dimensions $a = 9$ Å and $a = 10$ Å is less than 0.1 kJ/mol for KS/PBE, but surpasses 0.3 kJ/mol for KS/PBE0. The same behavior is observed when comparing the convergence of the RPA correlation energies between KS/SCAN and KS/SCAN0 orbitals. These results indicate that the converged

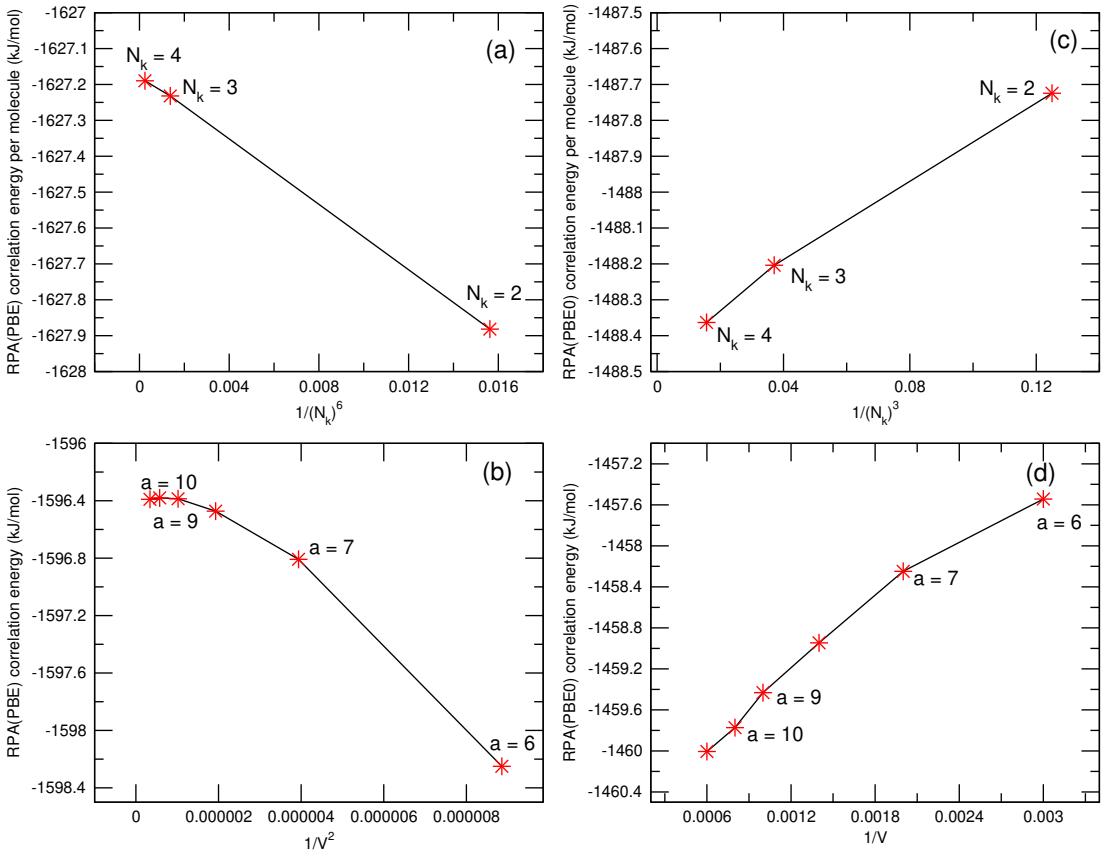

Figure 3.9: The convergence of the RPA correlation energy of ethylene solid and molecule with k-points and cell volume. RPA was evaluated on KS/PBE (a, b) and KS/PBE0 (c, d) states. A cut-off of 600 eV was used.

RPA values based on KS/PBE and KS/SCAN orbitals can be directly obtained with a sufficiently large number of k-points for solids or a sufficiently large cell volume for molecules, whereas extrapolations with these parameters are necessary to obtain converged RPA values for KS/PBE0 and KS/SCAN0 states (see Table 3.11). We note that the convergence of the RPA correlation energies based on hybrid functionals with k-points and cell volume can be improved when using the Coulomb cut-off technique at the step of DFT calculations.

Based on the convergence behavior of the RPA correlation energies discussed above, we can estimate the errors in the RPA correlation energies based on KS/PBE and KS/SCAN orbitals, caused by the convergence problem of k-points and cell volume, are below 0.1 kJ/mol for all the considered systems. However, for the RPA correlation energies based on KS/PBE0 and KS/SCAN0 orbitals, these errors are comparatively larger, falling within the range of 0.1 to 0.2 kJ/mol.

We now consider the convergence of the MP2 correlation energy. The computational cost for the calculations of the MP2 correlation energy (scales as $\mathcal{O}(N^5)$) is higher than that of the RPA correlation energy (scales as $\mathcal{O}(N^4)$). Obtaining the MP2 correlation energies for solids with a large number of k-points and a high plane-wave basis set is challenging due to high computational requirements. Based on the fact that the MP2 correlation energy can be decomposed into exchange contribution (xMP2) and direct contribution (dMP2), we only performed the xMP2 calculations using the standard MP2 implementation, while the dMP2

part was calculated using RPA algorithm starting from HF orbitals. It is clear that this way can reduce significantly computational cost for the MP2 calculations if the xMP2 component converges faster than dMP2. For the xMP2 part, we note that we could only perform a calculation up to $N_k = 3$ with a plane-wave basis-set cut-off (tag `ENCUT` in VASP) of 400 eV for ethane and ethylene, and up to $N_k = 2$ with a plane-wave basis-set cut-off of 700 eV for the two forms of acetylene, shown in Table 3.11.

For all the solids that we considered, the MP2 correlation energy is found to converge quickly with the number of k-points, similar to that of the RPA correlation energy based on KS/PBE and KS/SCAN orbitals. This rapid convergence is due to the use of the Coulomb cut-off technique in the HF states applied for the MP2 calculations. However, for the isolated molecules, we find that the MP2 correlation energy converges rather slowly with cell volume. Hence, in most cases, extrapolations are necessary, with only the exception being the dMP2 energy of ethane (see Table 3.11).

Based on the convergence behavior, the uncertainties associated with the convergence problems of k-points and cell volume for the MP2 correlation energies are estimated to be less than 0.2 kJ/mol for all the systems considered. The main source of uncertainty arises from errors in the xMP2 component. Additionally, the uncertainties are expected to be greater for the two forms of acetylene compared to ethane and ethylene due to the missing of xMP2 energy data for the former with large number of k-points (Table 3.11).

Finally, we note that it is difficult to obtain all the datapoints that are required to perform extrapolations as mentioned above. This is mainly due to the fact that the memory requirements for the RPA and MP2 calculations increase very fast with numerical parameters. To overcome this limitation, we adopted a strategy where we obtained the desired energy for a large cut-off and dense k-point grid ($E^{\text{dense}}_{\text{large}}$) by using known energies obtained with smaller cut-off and sparse k-point grid. This approach can be expressed as: $E^{\text{dense}}_{\text{large}} = E^{\text{sparse}}_{\text{large}} + (E^{\text{dense}}_{\text{small}} - E^{\text{sparse}}_{\text{small}})$ [21]. We applied the same strategy to isolated molecules, where energies were evaluated for different cell sizes instead of k-point grids.

3.2.2 Plane wave basis sets

We now consider the convergence of the energy components with plane-wave basis sets. We observe that the choice of DFT orbitals almost does not affect the convergence of the energy components in the RPA calculations with the basis sets. Moreover, the convergence is the same for all the considered systems. Therefore, we will discuss one system as a representative example.

We start to consider the convergence of the mean field energies with the basis set. Here, we show EXX of ethylene in Fig. 3.10 as an example. One can see that the energy difference in the EXX binding energy between ENCUT = 1000 eV and ENCUT = 1200 eV is very small (below 0.01 kJ/mol). Therefore, we used an ENCUT of 1100 eV to obtain converged values of all the mean field energies with the basis set. The exception here is the RSE component, where an ENCUT of 800 eV was used. The reason for this is that the convergence of the RSE with the basis set was found to be faster than the convergence of the other mean field energies. By using these cut-off values, the uncertainties

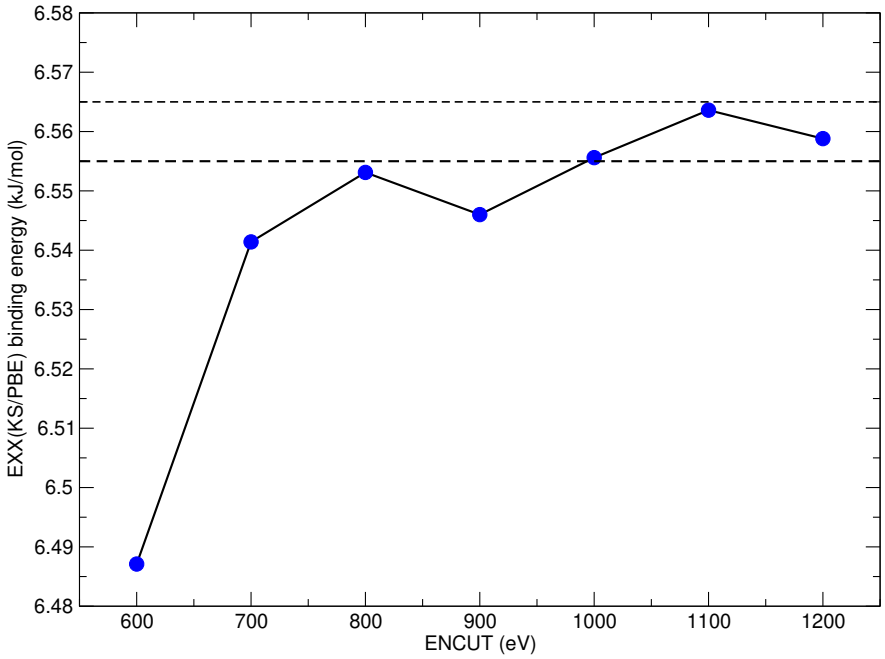

Figure 3.10: The convergence of the binding energy of EXX(KS/PBE) with the plane wave basis sets for ethylene. A k-points with $N_k = 2$ and a cell volume with $a = 8$ were used for the calculations.

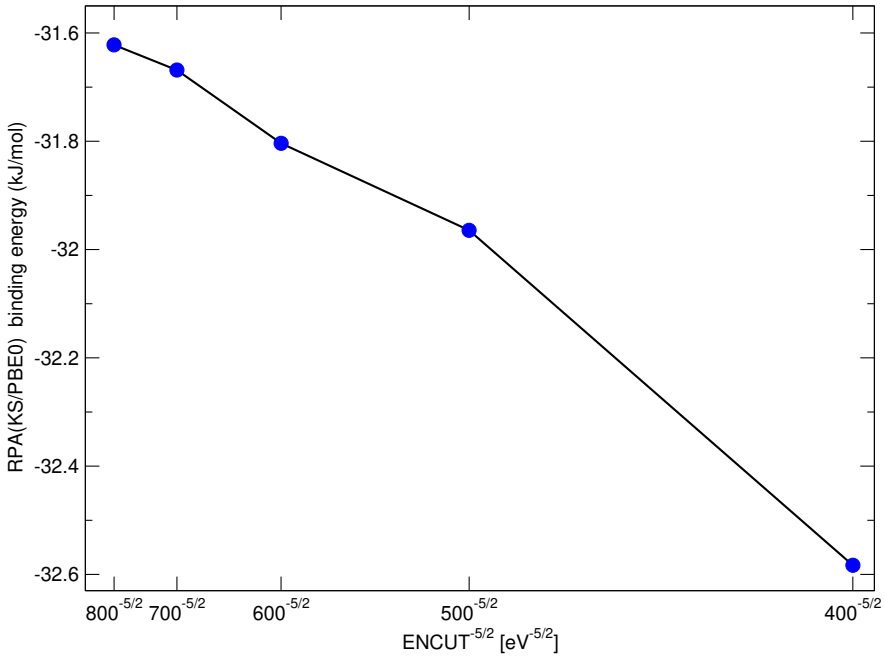

Figure 3.11: The convergence of the binding energy of RPA(KS/PBE0) with the plane wave basis sets for ethane.

from the basis set convergence problem for the mean field energies can be safely considered negligible.

We now move to discuss the convergence of the correlation energies (RPA and MP2) with the basis set size. These energy components converge slowly with the basis set, and extrapolations are used to obtain converged values. It was observed that the leading order error of the RPA correlation energy depends as $E_{cut}^{-3/2}$ on the plane-wave cut-off [275, 276]. However, for binding energy of molecular solids a different dependence being proportional to $E_{cut}^{-5/2}$ was found to dominate [21, 41]. We also used such convergence behavior for our calculations, shown in Fig. 3.11 as an example. In our calculations, the RPA and MP2 correlation energies were evaluated for different values of cut-off, from 400 eV to 800 eV with 100 eV increments. However, for the final extrapolations, we only considered cut-offs from 600 to 800 eV. This is due to numerical noise when the data for smaller cut-offs (400 and 500 eV) were used (see Fig. 3.11). This behavior was also observed in previous work [21]. The convergence behavior of the MP2 correlation energy with the basis set is similar to that of the RPA correlation energy.

To estimate the uncertainties of the correlation energies due to the basis set convergence problem, we have examined that the RPA binding energy obtained by extrapolating to infinite basis set using cutoffs from 600 to 1000 eV agrees within 0.01 kJ/mol with values extrapolated from 600 to 800 eV. This very small error indicates that using data within the range of 600 to 800 eV for extrapolations is sufficient to obtain converged binding energies of the RPA and MP2 correlation energies. Therefore, the uncertainties associated with the basis set convergence problem can be safely considered negligible for these energy components.

3.2.3 PAW potentials

As mentioned in the computational setup, the mean field energies (HF and RSE) and correlation energies (RPA and MP2) were evaluated with standard PAW potentials. To assess the uncertainties arising from the use of standard PAWs instead of hard ones, we took ethylene as an example and calculated the energy components using hard PAWs, employing the settings similar to those used for standard PAWs.

Table 3.12: The values of the binding energies of ethylene (in kJ/mol) obtained from standard and hard potentials.

Energy components	KS/PBE		KS/PBE0	
	Standard	Hard	Standard	Hard
RSE	−2.53	−2.57	−0.89	−0.92
EXX	12.66	12.63	10.60	10.57
RPA	−30.63	−30.68	−27.42	−27.53

The results presented in Table 3.9 show that the differences of the energy components between standard and hard potentials are small, being less than 0.15 kJ/mol. These errors should be considered when assessing the overall uncertainties. Notably, the errors in the EXX and RSE energies are smaller in comparison to those observed in the RPA component. The small effect of hard

potentials on the RSE energies was also observed previously [21]. Finally, we note that we used PAW potentials based on KS/PBE states for all the calculations. This choice may have an impact on the results of the energy components based on KS/SCAN and KS/SCAN0 states.

3.2.4 Summary

The numerical parameters affecting the results of the energy components in the RPA and MP2 calculations obtained from the PBC approach were discussed in detail. We find that achieving converged binding energy of a molecular solid with numerical parameters is challenging. The Coulomb cutoff technique [274] is beneficial for achieving the mean field energies converged with the k-points and cell volume. However, the computational and memory requirements pose limitations on achieving convergence of the RPA and MP2 correlation energies with respect to the k-points and cell volume. To obtain converged binding energy with the basis set, sufficiently large basis sets can be used for the mean field energies, while extrapolation to infinite basis-set limit is necessary for the RPA and MP2 correlation energies. Standard PAW potentials should be acceptable for obtaining energy components in the RPA and MP2 calculations for the considered systems if a precision loss of around 0.1 kJ/mol is acceptable.

4. Comparison between MBE and PBC results

Comparing the energy components obtained from the MBE approach with those obtained from PBC can provide valuable insights into the precision issues discussed in the previous chapter. An excellent agreement was found between MBE and PBC approaches for the energy components in the MP2 calculations [23]. However, the comparison of the energy components in the RPA calculations between these approaches has not yet been studied, and we will focus on discussing this in this chapter. We expect the differences in the energy components between MBE and PBC to be less than 0.1 kJ/mol. Any differences larger than 0.1 kJ/mol will be considered significant errors, likely arising from the convergence problem of the energy components with numerical parameters. Moreover, comparing the energy values in the RPA calculations obtained from the MBE and PBC approaches can also serve to validate the reliability of the RPA program applied in our MBE calculations. We will start to discuss the mean field energies, and then with the correlation energies.

4.1 Mean field energies

The mean field binding energies obtained from both MBE and PBC approaches are summarized in Tables A.1–A.4 in Attachments. Here we show the differences between the MBE and PBC values for these energy components in Fig. 4.1. We note that the differences between the MBE and PBC values are rather similar for EXX(KS/PBE) and EXX(KS/PBE0), as well as for EXX(KS/SCAN) and EXX(KS/SCAN0). For the RSE component, the differences are comparable when based on KS/PBE and KS/SCAN, and when based on KS/PBE0 and KS/SCAN0. Therefore, we only present EXX evaluated with KS/PBE and KS/SCAN, and RSE evaluated with KS/PBE and KS/PBE0 in Fig. 4.1 to make the graph simpler.

Let us start with the DFT and HF energies. One can see from Fig. 4.1 that the errors of DFT are less than 0.1 kJ/mol for KS/PBE and KS/PBE0 orbitals. These errors come from the slow convergence of the non-additive DFT contributions obtained from MBE with the cut-off distance (see Fig. 3.4(a) and 3.5(a) in chapter 3). However, the errors are notably larger for KS/SCAN and KS/SCAN0 orbitals. For example, the errors in DFT(KS/SCAN) are around 0.35 kJ/mol for ethane and ethylene. These large errors in the DFT energies based on KS/SCAN and KS/SCAN0 orbitals can be due to the numerical instabilities in the convergence of the DFT energies when using meta-GGA functionals [277]. Additionally, the use of PAW potentials based on KS/PBE can contribute to these large errors. Furthermore, they also originate from the slow convergence with the fragment size of the DFT energy compared to the other mean field energies. Indeed, the 4-body DFT values remain relatively large, reaching up to over 2 kJ/mol in some cases, whereas none of the other mean field energies predict such high values for the 4-body contributions (Tables A.2–A.4). Thus, the importance of the 5-body DFT contributions in the MBE calculations should

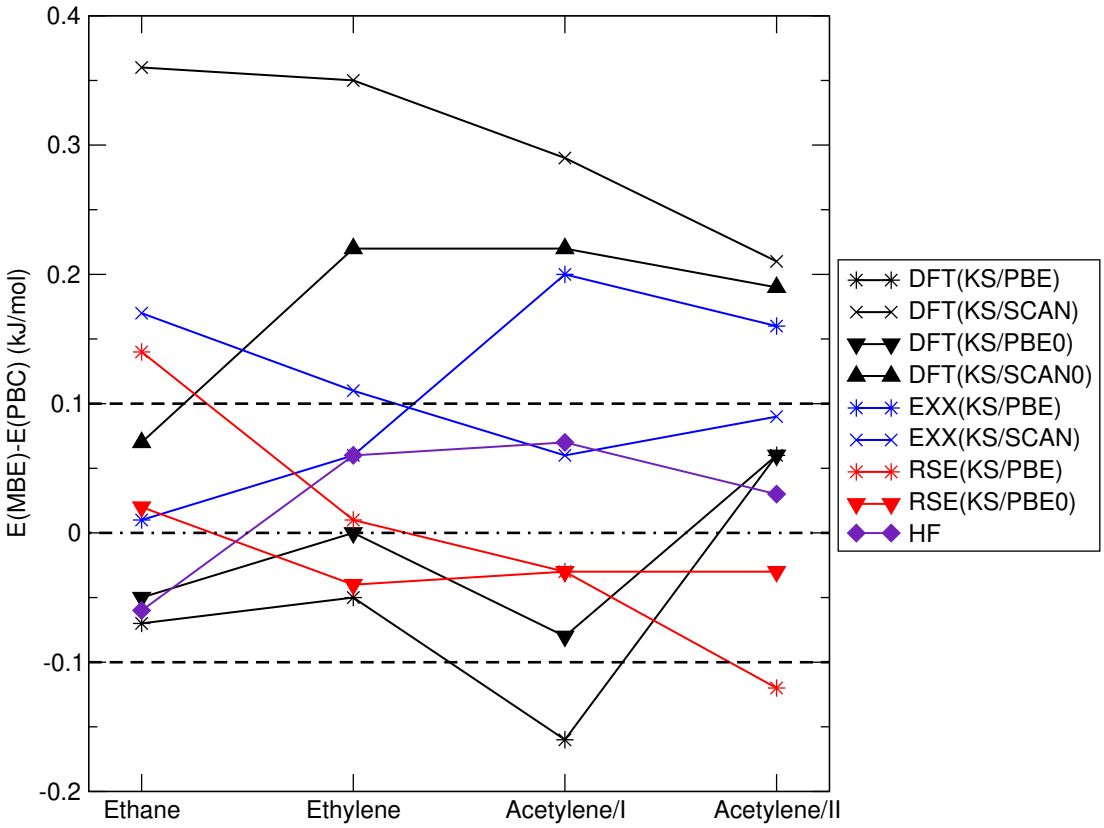

Figure 4.1: The differences in the mean field energies between MBC and PBC approaches.

be studied to better understand if the 5-body term would reduce the difference between MBE and PBC. The differences in the HF energy are very small (below 0.1 kJ/mol). These erros stem from the slow convergence of the non-additive HF contributions with the cut-off distance, similar to the DFT energies based on KS/PBE and KS/PBE0. The small errors for the HF energies were also observed for other systems [23].

The errors in the EXX energy are less than 0.2 kJ/mol (see Fig. 4.1). These errors are mainly due to the slow convergence of the non-additive EXX contributions with the cut-off distance in MBE (see Fig. 3.4(b) and 3.5(b) in chapter 3). For ethane and ethylene, the differences are larger for KS/SCAN and KS/SCAN0 than for KS/PBE and KS/PBE0. This may be again due to the use of PAWs based on KS/PBE in the PBC calculations. However, the opposite trend is observed for the two forms of acetylene. This observation can be a result of a quicker convergence with the fragment size of the EXX energy using KS/SCAN and KS/SCAN0 compared to KS/PBE and KS/PBE0. For instance, in the case of acetylene/I, the contribution of the 4-body EXX energy for KS/PBE is more than twice that for KS/SCAN (see Table A.3).

For the RSE component, the errors are below 0.15 kJ/mol. Similar to the other mean field energies, these errors in the RSE component are due to the convergence problem of the non-additive RSE contributions with the cut-off distance in MBE (see Fig. 3.5(c) in chapter 3). We note again that the RSE component in the PBC calculations was evaluated with standard PAWs. The variation in the RSE energy between standard and hard PAWs is small, being less than 0.05 kJ/mol

(see Table 3.9 in the previous chapter). However, this small variation can also contribute to overall errors in the RSE component. Compared to the DFT and EXX energies, the differences in the RSE energies between the MBE and PBC approaches are not significantly influenced by the orbitals studied. For all the systems considered, the differences are smaller for KS/PBE0 and KS/SCAN0 orbitals than for KS/PBE and KS/SCAN orbitals, with the exception of ethylene, where the opposite trend is observed. The smaller errors for KS/PBE0 and KS/SCAN0 compared to KS/PBE and KS/SCAN can result from a faster convergence of the RSE energy based on KS/PBE0 and KS/SCAN0 with the fragment size. For instance, in the case of the two forms of acetylene, the 4-body RSE contributions are nearly negligible when employing KS/PBE0, whereas they remain above 0.1 kJ/mol with KS/PBE (see Table A.4).

4.2 Correlation energies

We now turn to discuss the errors in the RPA and MP2 correlation energies, presented in Fig. 4.2. The binding energies of these correlation energies obtained from both MBE and PBC approaches are summarized in Table A.5 and Table A.6 in Attachments. We note that the 2-body contributions of the RPA and MP2 correlation energies obtained from the MBE approach were extrapolated to infinite cut-off distance assuming a r_{cut}^{-3} dependence of the binding energy on the cut-off distance r_{cut}. This extrapolation was done to enhance the reliability of the total RPA and MP2 correlation energies obtained from the MBE approach when compared to PBC.

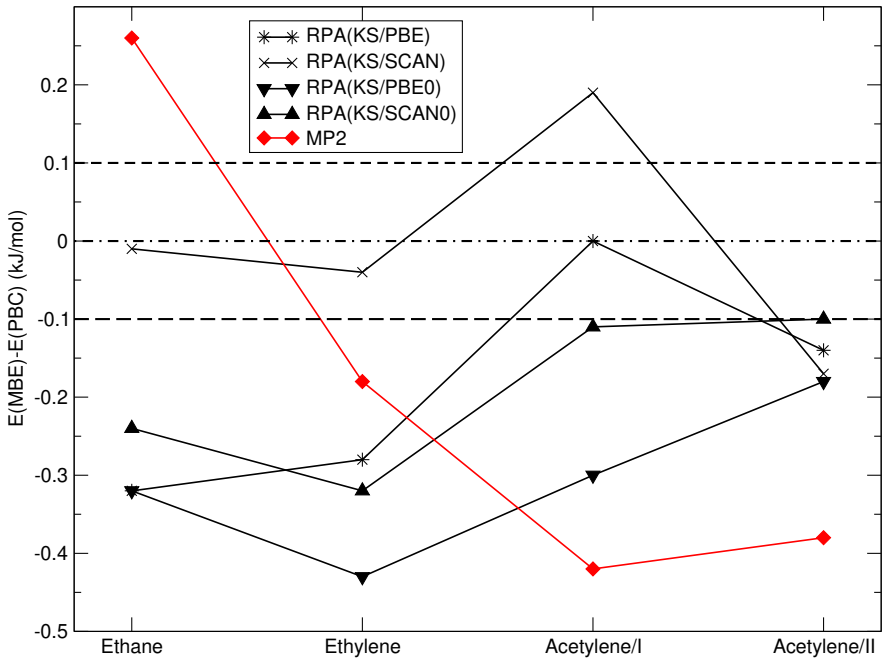

Figure 4.2: The differences in the correlation energies between MBC and PBC approaches.

The errors in the RPA correlation energies between two approaches are relatively large, but remain below 0.4 kJ/mol, i.e, around 2 %. As discussed in

the previous chapter, the RPA correlation energies evaluated with PBC converge slowly with numerical parameters, and extrapolations are necessary in some cases. These extrapolations can introduce large errors in the RPA correlation energies between two approaches. Additionally, the RPA correlation energies were obtained using standard PAWS, and the difference in the RPA correlation energies of around 0.1 kJ/mol between standard and hard PAWs (see Table 3.9 in chapter 3) can also contribute to these large errors. Finally, the slow convergence of the RPA correlation energies with the basis set and cut-off distance in MBE calculations can cause these significant errors. Indeed, the RPA correlation energies show a slower convergence with the basis set than the correlation energies based on WFT (see Table 3.1 as an example). The larger errors in the RPA correlation energies observed for hybrid orbitals compared to semilocal orbitals are mainly due to the slow convergence of the RPA correlation energies based on hybrid functionals with k-points or cell volume in the PBC calculations (see Fig. 3.9 in the previous chapter).

The MP2 correlation energies obtained from both approaches also exhibit acceptable agreement, with the errors below 0.4 kJ/mol. These errors are slightly larger than those observed for other systems [23]. However, they can stem from either several extrapolations of the MP2 correlation energies with parameters in PBC or from the slow distance convergence of the non-additive MP2 correlation energies in MBE, with the former being the dominant source of error. For example, the most significant differences, reaching approximately 0.4 kJ/mol, are observed in the case of the two acetylene forms, are mainly due to the need to use sparse k-point sets for the xMP2 energies for two these systems in PBC.

4.3 Summary

The differences between the MBE and PBC values are generally smaller for the mean field energies than for the correlation energies. The main errors in the mean field energies arise from the convergence problems related to the fragment size and cut-off distance in the MBE calculations. Thus, additional tests on the systems that exhibit a rapid convergence with fragment size and avoid the convergence issues with the cut-off distance can help to understand these errors. For correlation energies, the errors are primarily caused by the convergence problems with the numerical parameters in PBC compared to MBE. These errors can be significantly reduced when the numerical noise from extrapolations with k-points or cell volume, and basis-set cutoff, becomes negligible. The dependence of the correlation energies on k-points or cell volume can be reduced using the scheme developed by Liao and Grüneis [278]. Finally, the agreement between the MBE and PBC values for the energy components in the RPA calculations also shows reliability of the RPA program we used in our MBE calculations.

5. Effect of orbitals on energy components in RPA calculations

RPA calculations can be performed either non-self-consistently (based on DFT orbitals) or self-consistently. However, self-consistent RPA calculations are computationally demanding. Thus, most RPA calculations are currently based on DFT orbitals, in which KS/PBE orbitals are a common choice [160, 165, 279, 280, 281]. This is primarily due to their computational efficiency compared to meta-GGA or hybrid functionals. Moreover, RPA calculations based on KS/PBE orbitals have been widely used in solid-state studies [16, 21, 167, 282]. However, recent studies have shown that the quality of RPA results can be sensitive to the choice of orbitals [42, 43, 44, 282]. For example, for methane-water cage the variation in the RPA interaction energy between the KS/PBE and KS/SCAN orbitals was found to be around 1.5 kJ/mol [44]. This observation motivated us to evaluate the effect of different orbital inputs on the RPA results for molecular solids. As discussed in chapter 3, orbitals can affect significantly the basis set convergence of the RPA correlation energies obtained from the MBE approach. In this chapter, we will focus on understanding the influence of the orbitals on the results of the energy components in the RPA calculations. We find that the results are qualitatively similar for different systems. Therefore, we will focus on discussing ethylene as an illustrative example and refer to the results of other systems, which are summarized in Attachments, as necessary. The work presented here is composed of the results of our upcoming paper and paper P2.

5.1 Mean field energies

We start to discuss the mean field binding energies (DFT, EXX, HF, and RSE), focusing on how these energies change when they are evaluated on orbitals from hybrid functionals compared to semilocal functionals. One can see from Table 5.1 that the 2-body DFT energies based on all the examined functionals are attractive. The values derived from KS/SCAN and KS/SCAN0 exhibit stronger binding than those derived from KS/PBE and KS/PBE0. This behavior is consistent with observation in previous works [44, 89, 105]. The magnitude of the 2-body DFT interactions is observed to decrease by approximately 2 kJ/mol when transitioning from KS/PBE to KS/PBE0 and by around 0.8 kJ/mol when transitioning from KS/SCAN and KS/SCAN0 (Table 5.1). This shows that the 2-body DFT energies based on hybrid functionals tend to approach the 2-body HF energy, which is repulsive in the case of ethylene (see Table 5.1). This observation was also noted before in Ref. [44].

The 3-body DFT energies are repulsive for KS/PBE and KS/PBE0, while they are attractive for KS/SCAN and KS/SCAN0. These findings are also consistent with previous works [283, 284]. Conversely, the 4-body DFT energies exhibit an opposite trend compared to the 3-body DFT values for all the studied functionals. As shown in Table 5.1, the non-additive DFT energies based on KS/SCAN and KS/SCAN0 are closer to HF energy compared to those based on KS/PBE and KS/PBE0. Similar to the 2-body interactions, the effect of exact exchange on

Table 5.1: The mean field energies of ethylene in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE.

Components	Orbitals	2-body	3-body	4-body	Total
DFT	KS/PBE	−8.57	6.86	−1.86	−3.57
	KS/SCAN	−11.97	−3.08	1.48	−13.57
	KS/PBE0	−6.52	4.10	−1.19	−3.61
	KS/SCAN0	−11.18	−1.92	0.88	−12.24
EXX	KS/PBE	14.22	−2.15	0.62	12.69
	KS/SCAN	12.08	0.46	−0.31	12.23
	KS/PBE0	11.25	−0.79	0.19	10.65
	KS/SCAN0	10.52	0.33	−0.15	10.70
RSE	KS/PBE	−3.21	1.02	−0.33	−2.52
	KS/SCAN	−1.47	−0.60	0.16	−1.91
	KS/PBE0	−1.02	0.15	−0.06	−0.93
	KS/SCAN0	−0.34	−0.57	0.08	−0.83
1RDM,quad	semi/PBE	−1.50	0.72	–	–
HF	–	9.38	−0.38	0.02	9.02

the 3- and 4-body DFT energies is more significant when moving from KS/PBE to KS/PBE0 than from KS/SCAN to KS/SCAN0 (see Table 5.1).

The variations in n-body DFT contributions among the examined orbitals are quite large (Table 5.1). However, we find that the changes in n-body orders of DFT energy tend to cancel each other, leading to comparable total DFT binding energies for KS/PBE and KS/PBE0, as well as for KS/SCAN and KS/SCAN0. Finally, the differences in total DFT energies between semilocal functionals and their hybrid counterparts are found to be below 1.5 kJ/mol for all the considered systems.

We now turn to discuss the EXX component. The 2-body EXX energies based on all orbitals are more repulsive than the 2-body HF energy. Thus, the 2-body EXX energies become less repulsive when going from KS/PBE and KS/SCAN to KS/PBE0 and KS/SCAN0, respectively, which was also observed before in Ref. [44]. For the 3-body interactions, the addition of exact exchange in hybrids causes a decrease in the EXX binding energy for KS/PBE but an increase for KS/SCAN. The behavior observed for the 4-body terms is opposite to that of the 3-body terms. These changes in the 3-body and 4-body EXX energies occur because the values derived from hybrid functionals tend to approach the HF values (see Table 5.1). Similar to the behavior observed in the DFT energy, the changes in all n-body EXX energies when exact exchange is introduced are more significant for KS/PBE compared to KS/SCAN. However, these changes are smaller in magnitude than those observed for the DFT energy (see Table 5.1). Finally, we also find that there are cancellations in the errors of n-body EXX contributions, leading to differences of below 1.0 kJ/mol in total EXX binding energies between KS/PBE and KS/SCAN. The same behavior is observed for KS/PBE0 and KS/SCAN0.

Let us shift our discussion to the RSE component. The difference between

the EXX and HF values is expected to decrease when the RSE corrections are added to EXX energy. As shown in Table 5.1, adding exact exchange in hybrid functionals results in a decrease in the binding of the 2-body RSE energies based on both KS/PBE and KS/SCAN. This can be explained by noting that the 2-body EXX energies based on hybrid functionals are closer to the 2-body HF energy than those based on semilocal functionals. Therefore, the strength of the 2-body RSE interactions is smaller for hybrids than for semilocal functionals. Transitioning from KS/PBE to KS/PBE0 results in a decrease in the repulsive interaction within the 3-body RSE energies and the attractive interaction within the 4-body RSE energies. These trends can also be understood in the context of the changes observed in the EXX energies, as discussed above for the 2-body terms. Remarkably, there is almost no change in the 3-body and 4-body RSE energies when going from KS/SCAN to KS/SCAN0 (see Table 5.1 for ethylene and Tables A.7–A.9 for the other systems). This means that exact exchange has a minimal effect on the non-additive RSE corrections based on KS/SCAN.

5.2 Correlation energies

We now discuss how the correlation energies are influenced by the input states, starting with the RPA component. One can see from Table 5.2 that the 2-body RPA energies are more attractive when using KS/PBE and KS/SCAN orbitals compared to other orbitals. This difference originates from smaller electronic gap observed for RPA based on semilocal DFT orbitals than for RPA based on other orbitals, which was also noted before for other systems [43, 44]. Similar to the behavior observed for the mean field energies, the change in the 2-body RPA correlation energy from KS/PBE to KS/PBE0 is larger than the change from KS/SCAN to KS/SCAN0. As discussed in chapter 2, semi/PBE orbitals are an intermediate choice between KS/PBE and HF orbitals. Thus, the 2-body RPA(semi/PBE) correlation energy is found to be less attractive than 2-body RPA(KS/PBE) but more attractive than 2-body RPA(HF) for ethane (see Table A.10). However, for the systems with π electrons, the 2-body RPA(semi/PBE) energy is notably less attractive than 2-body RPA(KS/PBE), but relatively close to the HF values. For example, for the two forms of acetylene, the difference between the 2-body RPA(semi/PBE) and 2-body RPA(HF) energies is very small (below 0.15 kJ/mol) (see Tables A.11 and A.12). These observations indicate that semi/PBE orbitals can be used to replace HF orbitals for performing the 2-body RPA calculations in the systems with high electrostatic contributions. We note that the reduced binding of the RPA(semi/PBE) correlation energy is then corrected by the 2g term when both these components contribute to the RPA+MBPT3(semi/PBE) method.

The 3-body RPA correlation energies based on all examined orbitals are repulsive. When going from KS/PBE to KS/PBE0, the repulsive interaction in the 3-body RPA values decreases for ethane and ethylene (see Table A.10 and Table 5.2). This trend aligns with the change in the 3-body DFT energy from KS/PBE to KS/PBE0, and has been observed previously in Ref. [44]. However, we observe an increase in the 3-body RPA repulsive term for the two forms of acetylene when moving from KS/PBE to KS/PBE0, as shown in Tables A.11 and A.12. When going from KS/SCAN to KS/SCAN0, the changes in the 3-body

Table 5.2: The correlation energies of ethylene in kJ/mol obtained from the RPA calculations using the MBE approach.

Components	Orbitals	2-body	3-body	4-body	Total
RPA	KS/PBE	−32.15	1.07	0.24	−30.84
	KS/SCAN	−30.91	0.47	−0.02	−30.46
	KS/PBE0	−28.90	1.01	0.09	−27.80
	KS/SCAN0	−28.44	0.58	−0.01	−27.87
	HF	−24.12	1.49	−0.11	−22.74
	semi/PBE	−25.31	0.66	–	–
SOSEX	KS/PBE	−2.31	0.69	–	–
	semi/PBE	−0.74	0.41	–	–
	HF	0.11	0.03	–	–
2g	semi/PBE	−5.11	0.34	–	–
	HF	−5.22	0.61	–	–

RPA correlation energy are minimal (below 0.15 kJ/mol) for all considered systems. This is also consistent with the changes in the 3-body DFT energy, where the shift from KS/SCAN to KS/SCAN0 is significantly smaller than the change from KS/PBE to KS/PBE0, as discussed above. The 3-body RPA energies derived from HF orbitals (approximately 1-2 kJ/mol) exhibit significantly stronger repulsion compared to those computed from DFT orbitals, which aligns with the differences between the 3-body HF and 3-body DFT energies. The 3-body RPA(semi/PBE) values are less repulsive than both the 3-body RPA(KS/PBE) and RPA(HF) energies for all the considered systems.

Adding exact exchange into KS/PBE0 results in a reduction of the repulsive interactions observed in the 4-body RPA(KS/PBE) values. An exception to this trend is noted in the case of ethane, but the change is very small. As seen in Table 5.2, the exact exchange has almost no impact on the 4-body RPA value based on KS/SCAN orbitals for ethylene. The same behavior is observed for other systems (Tables A.10–A.12).

Similar to the behavior seen in the mean field energies, the variations in n-body RPA correlation energies tend to cancel each other. As a result, the total RPA correlation energies based on KS/PBE and KS/SCAN orbitals, as well as KS/PBE0 and KS/SCAN0 orbitals, are comparable.

We now move to discuss the SOSEX and 2g corrections. We note again that these components were obtained only up to the 3-body terms. The attractive interaction in the 2-body SOSEX corrections is significantly decreased when going from KS/PBE to HF orbitals. This is expected as the SOSEX corrections are added to remove second-order Pauli-exclusion violating contributions in the RPA correlation energy using DFT orbitals. As discussed above, the 3-body RPA correlation energies based on HF orbitals are much more repulsive than those based on DFT orbitals. Consequently, the repulsive interactions in the 3-body SOSEX energies are found to decrease when going from KS/PBE to HF orbitals. The 2g values based on both semi/PBE and HF orbitals are attractive, with only a minimal difference of 0.1 kJ/mol. However, the magnitude of the repulsive 3-body 2g values obtained from HF orbitals is double of that obtained

for semi/PBE orbitals. These differences are consistent with differences between the RPA(semi/PBE) and RPA(HF) correlation energies, discussed above.

5.3 Summary

The results of the energy components in the RPA calculations are affected by the choice of orbital inputs. Adding exact exchange in hybrid functionals has a greater effect on the change in energy of both the mean field and RPA correlation energies when based on KS/PBE orbitals compared to KS/SCAN orbitals. Remarkably, there is almost no change in the non-additive energies of RSE and RPA correlation energies when KS/SCAN input is replaced by KS/SCAN0. For all the energy components, the changes in n-body contributions between KS/PBE and KS/SCAN orbitals are relatively large, but the errors in n-body contributions tend to cancel each other, resulting in relatively small differences in the total energies between these orbitals. The same behavior is observed for KS/PBE0 and KS/SCAN0 orbitals. However, there are exceptions regarding the total DFT energies, where the values derived from KS/PBE are similar to those from KS/SCAN, and similarly, the values from KS/PBE0 are comparable to those from KS/SCAN0.

6. Accuracy of examined approximate methods

For molecular solids, the accuracy of approximate methods is commonly assessed by comparing their binding energies to reference data [16, 21, 28, 285]. However, this approach has limitations as a single binding energy value is often insufficient to fully understand the deviations of approximate methods from the reference. To obtain a more comprehensive understanding, it is preferable to evaluate each n-body contribution of the binding energy obtained from MBE by comparing it with the reference [31]. In this chapter, we first discuss the results which were published in papers P1 and P2 where we tested the accuracy of MP2, MP3, RPA and beyond-RPA methods in describing of n-body terms contributing to the binding energy of all the considered systems. In the remaining part of this chapter, we will examine the performance of the correction scheme, where the n-body contributions at large distance obtained from the CCSD(T) reference are substituted by those obtained from approximated methods. Then, we evaluate the accuracy of some dispersion-corrected DFT methods obtained from the PBC approach. These results will be presented in our forthcoming work. As the results are spread over different papers, we combine them here to give a broader perspective.

6.1 Two-body terms

The 2-body contributions of the examined methods for all the considered systems are summarized in Table A.13 in Attachments, and are compared with reference CCSD(T) data in Fig. 6.1. Let us start to consider the RPA method without any additional corrections. Fig. 6.1 shows that the total 2-body RPA energies based on all studied orbitals significantly underestimate the reference CCSD(T) data. The average relative differences calculated for all the considered systems compared to the reference data range from approximately 18% to 32%. Specifically, the RPA(KS/SCAN) values exhibit the closest agreement with the reference data, while RPA(HF) values give the largest errors. This observation was also found previously for other system [44]. The reason for this is that the RPA correlation energies based on HF orbitals are considerably less attractive than those based on DFT orbitals, due to the larger electronic gap of HF.

The addition of the RSE corrections improves significantly the accuracy of the direct RPA method, in which the values based on pure DFT functionals (KS/PBE and KS/SCAN) show a better agreement with the reference than those based on hybrids (KS/PBE0 and KS/SCAN0). RPA+RSE shows better performance for KS/PBE than for KS/SCAN, primarily due to the more significant influence of the RSE corrections based on KS/PBE. These findings are also consistent with those reported in Ref. [44].

We now turn to examine the RPA methods with the remaining corrections beyond RSE. The SOSEX corrections are applied to remove the second-order Pauli-exclusion violating contributions in the RPA correlation energy evaluated with DFT orbitals, and we find that their contributions calculated with KS/PBE

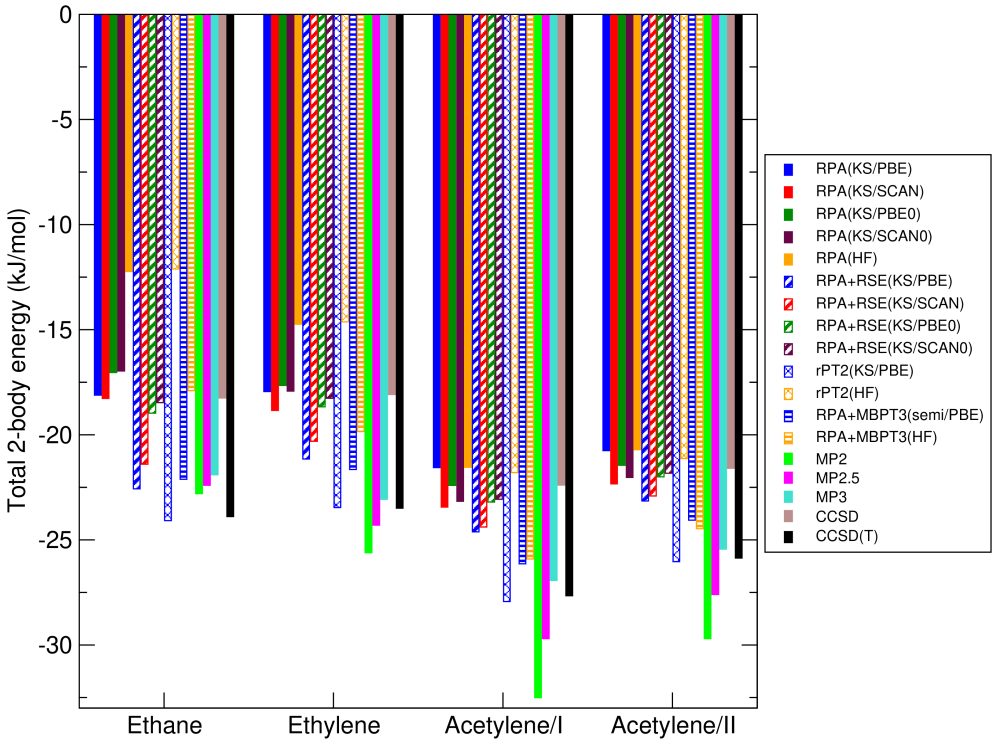

Figure 6.1: The two-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data.

orbitals are important for the 2-body terms of the considered systems. For instance, the SOSEX(KS/PBE) contribution is even larger than the contribution of RSE(KS/PBE) for acetylene/I (refer to Table A.8 and Table A.11). RPA+RSE(KS/PBE) is improved significantly when the SOSEX(KS/PBE) corrections are added. Indeed, the agreement of the rPT2(KS/PBE) method, where both RSE and SOSEX corrections are added to total RPA, with the CCSD(T) reference is excellent and the relative errors are below 1% for all the considered systems (see Fig. 6.1). We note that this exceptional agreement arises not from error cancellation, but from the observation at the level of each individual dimer in the 2-body contributions, see Fig. 6.2 as an example. The improvement of rPT2(KS/PBE) over RPA+RSE(KS/PBE) was also noted previously for dispersion-dominated systems [181]. However, it is important to note that rPT2(KS/PBE) yields significant errors for hydrogen-bonded systems [181]. Therefore, the accuracy of the rPT2(KS/PBE) method should be examined for molecular solids with hydrogen bonds.

The RPA+MBPT3 method was evaluated with addition of all corrections considered (RSE, 1RDM,quad, SOSEX, and 2g). We note again that the EXX and RSE values do not change when going from KS/PBE to semi/PBE orbitals. The 2-body RPA and SOSEX correlation energies evaluated with semi/PBE are less attractive than when evaluated with KS/PBE. The diminished attraction of the RPA(semi/PBE) correlation energies can be improved by the 2g corrections, but it is not fully compensated. Consequently, when adding all terms together, RPA+MBPT3(semi/PBE) offers only a slight improvement over RPA+RSE(KS/PBE). Their relative errors compared to the CCSD(T) reference data are still larger than 5% for all the considered systems. The general under-

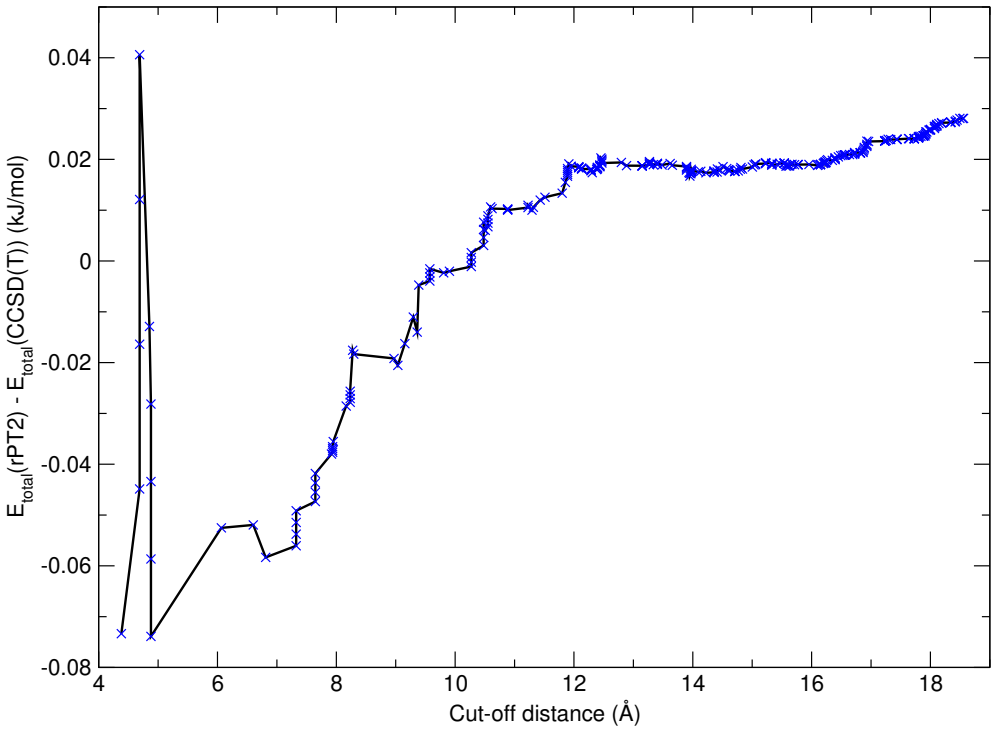

Figure 6.2: The difference between the rPT2(KS/PBE) and CCSD(T) binding energies over distance for the 2-body contributions of ethylene.

binding trend of RPA+MBPT3(semi/PBE) for high-dispersion systems was also observed before in Ref. [45].

When using HF orbitals as a starting point, the contributions of the 2-body SOSEX corrections are very small (see Table 5.2 in the previous chapter as an example). This leads to a similar level of accuracy between the RPA(HF) and rPT2(HF) methods, as shown in Fig. 6.1. However, a significant improvement is observed with RPA+MBPT3(HF) compared to RPA(HF). Remarkably, for the two forms of acetylene, the precision achieved with RPA+MBPT3(HF) is almost comparable to that of RPA+MBPT3(semi/PBE). The notable improvement of RPA+MBPT3(HF) over RPA(HF) is a consequence of the effective compensation of the 2g(HF) corrections for the reduced attractiveness of the RPA(HF) correlation energies.

We now shift our focus to the examination of the WFT methods. The CCSD energies show an average error for all the systems from the reference data of around 20 %, with larger errors observed for ethane and ethylene compared to the two forms of acetylene. These large errors are due to the absence of the 2-body (T) component, which plays an important role for the 2-body energies in general. For MP2, we find that the deviation from CCSD(T) increases notably when going from ethane to the systems with delocalized electrons. The relative errors range from around 4.6% for ethane to as high as 17.5% for acetylene/I. This result can be understood by the limitation of MP2 in describing the systems with delocalized electrons [286, 287]. However, MP3 exhibits an opposite trend. Specifically, the 2-body MP3 energies deviate from the reference data by less than 3% for the systems with π electrons, but reach up to 8% for ethane. These errors show that the MP3 method outperforms all the examined methods, apart from

rPT2(KS/PBE), for the systems with delocalized electrons. Finally, the MP2.5 method performs quite well for all the considered systems, generally exhibiting smaller errors compared to those observed in the RPA and beyond-RPA methods based on HF orbitals. These results indicate that the MP2.5 and MP3 methods are better choices compared to the beyond-RPA(HF) methods for the 2-body contributions of the considered systems, at least in terms of accuracy.

Understanding the accuracy of the approximate methods for proximate and distant dimers separately is important as the contributions of the energy components depend strongly on the cut-off distance. For the proximate dimers, we find that the errors of the examined methods compared to the CCSD(T) reference data are mostly similar to those for the dimers summed over all distance. For example, rPT2(KS/PBE) also yields the errors of less than 1% for all the considered systems when examining the proximate dimers. This similarity arises because the largest contribution of the energy components typically comes from short intermolecular separations. Therefore, we only discuss the accuracy of the approximate methods for the distant dimers, summarized in Table 6.1. The interactions at large distances are governed by the leading terms which decay as $(-C_6/R^6)$ with the intermolecular distance R. By examining the distant dimers, we can assess how accurately approximate methods predict the leading order of the interaction energy. Moreover, we can test how the leading order is affected by the choices made for RPA orbital inputs or by the addition of further MP or CC terms.

We first consider the performance of the RPA and beyond-RPA methods for the distant dimers. When compared to the CCSD(T) reference data, the direct RPA method yields errors of less than 0.2 kJ/mol, with RPA(KS/PBE) exhibiting the smallest errors (see Table 6.1). Then there is almost no improvement when going from direct RPA(KS/PBE) to RPA(KS/PBE) with additional corrections. This is due to the fact that the RSE(KS/PBE) and SOSEX(KS/PBE) corrections converge rapidly with distance, and their contributions become nearly negligible at large distances. The 2g corrections account for third-order RPA corrections and are known to converge rather slowly with distance, as discussed before in chapter 3. The contributions of the 2g(semi/PBE) corrections are approximately 0.1 kJ/mol for all the considered systems. However, the contribution of the RPA(semi/PBE) correlation energies for the distant dimers are smaller compared to that of RPA(KS/PBE) (see Table 3.7 in chapter 3). Consequently, the errors in RPA+MBPT3(semi/PBE) are also similar to those observed for RPA(KS/PBE). These observations imply the direct RPA(KS/PBE) method gives the results close to CCSD(T) for the distant dimers with the errors below 0.1 kJ/mol.

We now move to discus the WFT methods. Among these methods, MP2.5 and MP3 outperform MP2 and CCSD for the distant dimers. The errors of the MP2.5 and MP3 methods compared to the CCSD(T) reference are below 0.05 kJ/mol for all the considered systems. These errors are even smaller than those observed for the RPA and RPA-beyond methods. The exception here is ethane, where MP2 shows an excellent agreement with the reference data. However, the interactions are clearly too large for the systems with π electrons.

In summary, the RPA methods based on KS/PBE orbitals perform better than those based on HF and other DFT orbitals. The direct RPA methods without corrections perform poorly for the 2-body contributions of all the considered

Table 6.1: Total 2-body energies of the examined methods for distant dimers ($10 < r < r_{\text{cut}}$) (kJ/mol) for all the considered systems.

Methods	Ethane	Ethylene	Acetylene/I	Acetylene/II
		KS/PBE		
$E_{\text{total}}^{\text{RPA}}$	−0.54	−0.45	−0.53	−0.22
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.54	−0.45	−0.48	−0.24
$E_{\text{total}}^{\text{rPT2}}$	−0.57	−0.48	−0.51	−0.23
		Semi/PBE		
$E_{\text{total}}^{\text{RPA+MBPT3}}$	−0.53	−0.44	−0.41	−0.24
		KS/SCAN		
$E_{\text{total}}^{\text{RPA}}$	−0.50	−0.40	−0.43	−0.23
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.50	−0.40	−0.43	−0.23
		KS/PBE0		
$E_{\text{total}}^{\text{RPA}}$	−0.48	−0.39	−0.42	−0.22
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.48	−0.39	−0.42	−0.21
		KS/SCAN0		
$E_{\text{total}}^{\text{RPA}}$	−0.46	−0.37	−0.41	−0.20
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.46	−0.37	−0.41	−0.20
		HF		
$E_{\text{total}}^{\text{RPA}}$	−0.37	−0.31	−0.38	−0.16
$E_{\text{total}}^{\text{rPT2}}$	−0.39	−0.32	−0.38	−0.15
$E_{\text{total}}^{\text{RPA+MBPT3}}$	−0.48	−0.41	−0.45	−0.22
$E_{\text{total}}^{\text{MP2}}$	−0.61	−0.56	−0.63	−0.42
$E_{\text{total}}^{\text{MP2.5}}$	−0.59	−0.53	−0.55	−0.33
$E_{\text{total}}^{\text{MP3}}$	−0.57	−0.49	−0.48	−0.25
$E_{\text{total}}^{\text{CCSD}}$	−0.51	−0.41	−0.43	−0.22
$E_{\text{total}}^{\text{CCSD(T)}}$	−0.61	−0.51	−0.52	−0.30

systems, but the inclusion of the additional corrections significantly improve their accuracy. Among the RPA variants, rPT2(KS/PBE) achieves the best accuracy for all considered systems. For the wavefunction methods, MP2 is suitable only for ethane, while MP3 is a good choice for the systems with delocalized electrons. The differences in the 2-body energies between approximate methods and CCSD(T) reference are notably smaller for the distant dimers compared to the proximate dimers. Therefore, one can consider using the 2-body energies of the distant dimers obtained from approximate methods, such as the RPA methods based on KS/PBE or MP2.5 and MP3 methods, instead of the CCSD(T) reference without sacrificing significant accuracy, thereby saving computational costs.

6.2 Three-body terms

We now turn to discuss the 3-body terms. The data for the examined methods are summarized in Table A.14, and are compared to the reference CCSD(T) data in Fig. 6.3. To begin, we focus on the RPA method without any additional corrections. According to Fig. 6.3, RPA based on KS/SCAN and KS/SCAN0 per-

forms much better than that based on KS/PBE and KS/PBE0. Remarkably, the RPA(KS/PBE) energies even have incorrect signs for all the considered systems. These observations are consistent with those observed in Ref. [44]. The accuracy of the RPA(HF) method is similar to the RPA(KS/SCAN) and RPA(KS/SCAN0) methods for ethane, aligning with findings reported for methane clathrate [44]. This can be clarified by noting that the EXX energies using KS/SCAN and KS/SCAN0 are close to the HF energies than when KS/PBE and KS/PBE0 are used. However, for the remaining systems, RPA based on HF orbitals exhibits superior performance compared to that based on DFT functionals. This indicates that HF orbitals are more suitable than DFT orbitals for describing the 3-body RPA energies for the systems with π interactions.

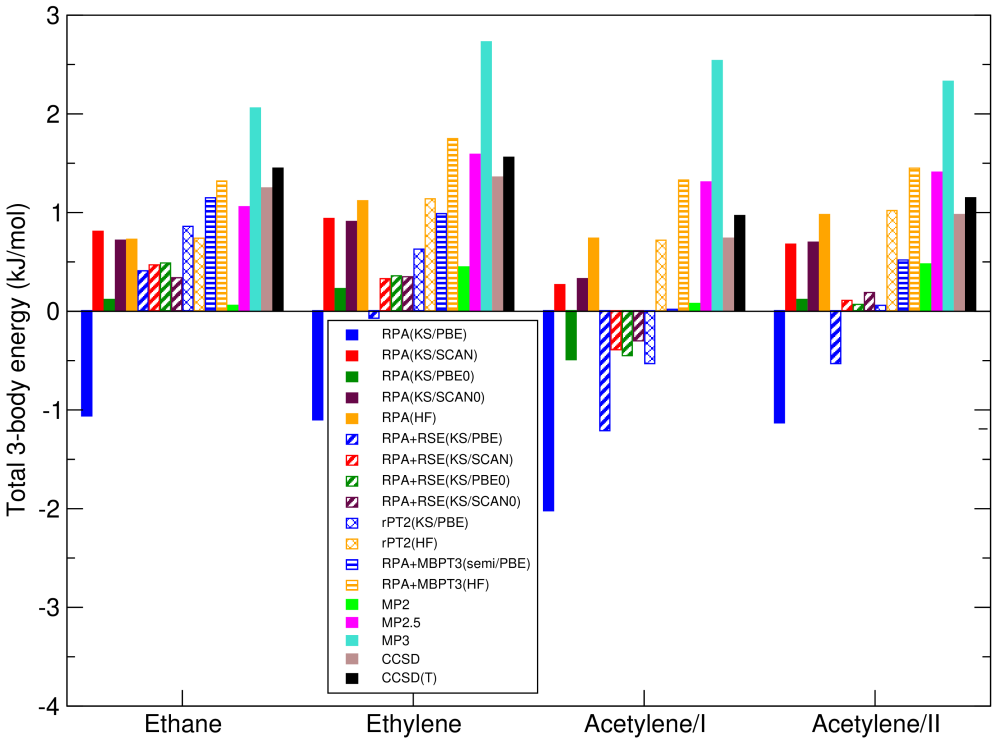

Figure 6.3: The three-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data.

We now discuss the results obtained for RPA with the RSE corrections added. The effect of these corrections varies depending on the type of DFT orbitals used. For KS/PBE and KS/PBE0, the 3-body RSE energies are positive and cancel partly the negative values of the 3-body EXX energies, and therefore they enhance the accuracy of RPA. By contrast, the RSE corrections reduce the accuracy of RPA when using KS/SCAN and KS/SCAN0. This behavior was also noted before and is possibly due to over-correction of the RSE corrections in the errors of the 3-body energies [43]. This recommends not to use the RSE corrections with the 3-body RPA(KS/SCAN) and RPA(KS/SCAN0) energies. The effect of the RSE corrections is notably more pronounced for KS/PBE orbitals than for KS/PBE0 orbitals. The average improvement is approximately 1 kJ/mol for the former but only 0.1 kJ/mol for the latter, as expected due to smaller RSE corrections for hybrid functionals. Thus, the RSE corrections may be not necessary for the 3-body RPA(KS/PBE0) energies, especially for the systems with delocalized

electrons.

Let us move to discuss the RPA methods with additional corrections beyond RSE. While the SOSEX corrections based on KS/PBE and semi/PBE were deemed negligible for the 3-body energy of methane clathrate [45], they play a significant role in the 3-body contributions of the considered systems. For instance, the SOSEX(KS/PBE) energy amounts to more than a half of the RSE energy for ethylene (see Tables 5.1 and 5.2). When combining the positive values of the SOSEX(KS/PBE) and RPA(KS/PBE) correlation energies, the resulting RPA+SOSEX(KS/PBE) values become closer to the CCSD(T) correlation energies. This makes the rPT2(KS/PBE) method much more accurate than the RPA+RSE(KS/PBE) variant. The accuracy of rPT2(KS/PBE) is then improved further by RPA+MBPT3(semi/PBE). Specifically, the average error is reduced by around 0.6 kJ/mol when using RPA+MBPT3(semi/PBE) compared to rPT2(KS/PBE).

When HF orbitals are used, the rPT2(HF) method demonstrates a comparable level of accuracy to RPA(HF), as the contribution of the 3-body SOSEX(HF) energies is minimal (below 0.03 kJ/mol) for all the systems considered. The RPA+MBPT3(HF) method is more accurate than rPT2(HF) for ethane and ethylene, whereas it is slightly inferior to rPT2(HF) for the two forms of acetylene. This discrepancy can be explained by the fact that the rPT2(HF) energies are much closer to CCSD(T) for the two forms of acetylene than for ethane and ethylene. Then the addition of the 2g corrections can improve the accuracy of rPT2(HF) for ethane and ethylene but may lead to overcorrection for the two forms of acetylene. Hence, when using HF orbitals for the two forms of acetylene, additional corrections may be unnecessary as the basic RPA method is sufficiently good.

We now examine the accuracy of the WFT methods. Among these methods, CCSD yields the 3-body energies that are in the best agreement with the reference for all the considered systems due to small contributions of the 3-body (T) component. The MP2.5 method is also an excellent choice for the 3-body terms. The errors of MP2.5 are slightly smaller than those of RPA+MBPT3(HF) for all the considered systems, except for ethane, for which the error of MP2.5 is around 0.4 kJ/mol. This is a consequence of error cancellation between MP2 (underestimation) and the MP3 (overestimation) binding energies, which was also noted before for the 3-body contributions of molecules [288]. The errors in the MP2 and MP3 methods are quite large, ranging from 0.5 to 1.6 kJ/mol. Specifically, MP2 performs better for the two forms of acetylene than for ethane and ethylene, while the opposite trend is observed for MP3. These trends are in contrast to the behavior observed in the 2-body terms.

As with the 2-body terms, the performance of the approximate methods for the 3-body terms is also influenced by distance. Thus, it is important to examine the performance of those methods as a function of the separation of molecules in a cluster. To do this, we divided all trimers into four groups based on the number of close contacts between molecules ($n = 0, 1, 2, 3$). Two molecules are in contact when $R < 5$ Å for ethane and acetylene/I and $R < 6$ Å for ethylene and acetylene/II, where R is intermolecular distance between molecules. We then analyze the errors of the approximate methods as a function of the molecules in contact, as shown in Fig. 6.4.

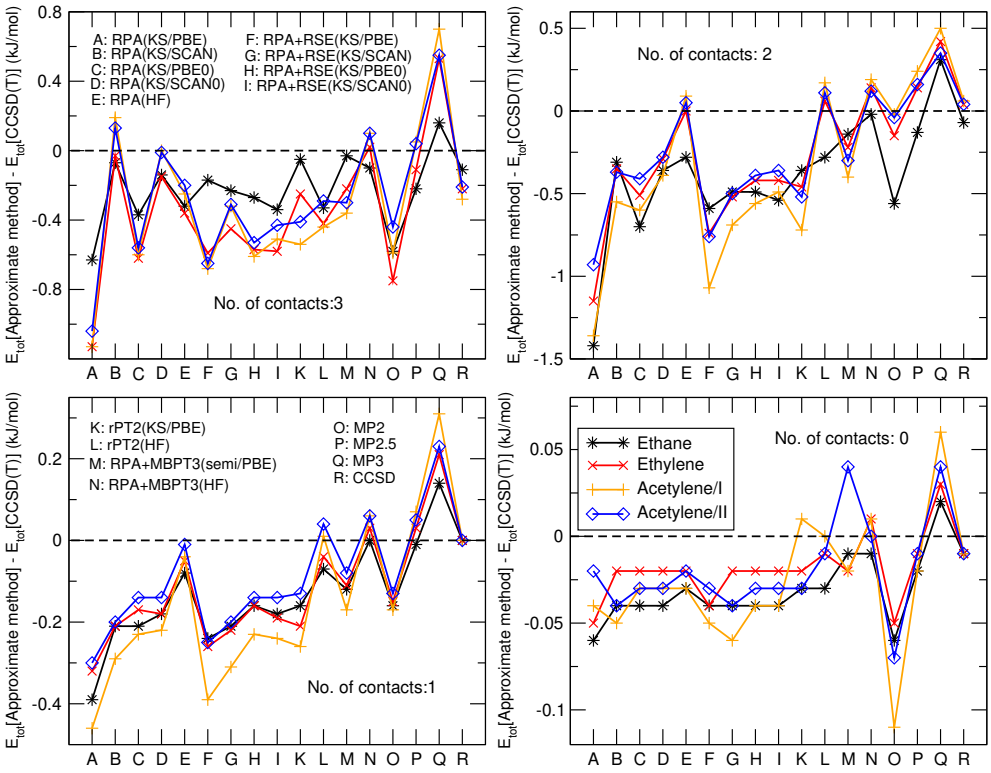

Figure 6.4: Signed errors in the three-body nonadditive energies of all the considered systems. The results are grouped according to the number of close contacts between the molecules in a trimer. Note the different scales on y axis.

We start to examine the RPA methods. One can see from Fig. 6.4 that at long range, for the systems with no close contacts ($n = 0$), the RPA methods without any corrections based on all studied orbitals yield the errors of below 0.05 kJ/mol compared to the CCSD(T) reference. However, we note that the CCSD(T) energies for $n = 0$ are rather small (between 0.1 and 0.3 kJ/mol) for the considered systems. Therefore, the errors of around 0.05 kJ/mol are small in magnitude, but relatively large. The inclusion of additional corrections only changes the errors of the RPA methods marginally.

For the groups with close contacts ($n = 1, 2, 3$), the 3-body RPA energies show considerable errors from the CCSD(T) reference. These errors decrease significantly when the RSE corrections based on KS/PBE and KS/PBE0 orbitals are added, but increase when including the RSE corrections based on KS/SCAN and KS/SCAN0 orbitals. This is in agreement to the behavior observed for the total 3-body energies. The largest errors in the RPA+RSE methods based on DFT orbitals are primarily observed in the group with two contacts, whereas for the RPA(HF) method, the most significant errors originate from the group with three contacts. This may result from the slower convergence of the RPA correlation energies based on DFT orbitals with distance compared to those based on HF orbitals (see Fig. 3.2).

The remaining additional corrections also influence significantly the performance of the approximate methods for the groups with close contacts. As shown in Fig. 6.4, rPT2(KS/PBE) is more accurate than RPA+RSE(KS/PBE), and RPA+MBPT3(semi/PBE) outperforms rPT2(KS/PBE) for all the groups with

close contacts. However, the errors in the RPA+MBPT3(semi/PBE) method compared to the CCSD(T) reference are still larger than 0.1 kJ/mol for the groups with close contacts. The only exception is the group with $n = 3$ of ethane, where RPA+MBPT3(semi/PBE) yields an error of 0.03 kJ/mol.

When based on HF orbitals, the RPA methods with and without corrections tend to perform better than those based on KS/PBE orbitals in trimers with close contacts between the molecules. For the group with $n = 1$, the RPA(HF) method yields very small errors (below 0.1 kJ/mol) for all the considered systems. The inclusion of additional corrections into RPA(HF) almost does not change these errors a lot. However, for the group with $n = 3$, RPA+MBPT3(HF) shows a significant improvement over RPA(HF). We observe this behavior for the group with $n = 2$ only for ethane. For the other systems, RPA+MBPT3(HF) performs slightly less effectively than RPA(HF). These observations suggest that the 3-body RPA(HF) energies can be used to replace the CCSD(T) reference for the groups of $n = 1, 2$, but for the group of $n = 3$, the RPA+MBPT3(HF) results are the preferable choice.

We now move to examine the WFT methods. The CCSD method shows substantial differences with the reference CCSD(T) only for the group with three contacts, ranging from 0.1 to 0.3 kJ/mol. For the remaining groups, it yields relatively accurate results, with errors below 0.1 kJ/mol. Nevertheless, it is worth noting that the CCSD method is computationally expensive. The other methods (MP2, MP2.5, and MP3) show significant differences with the reference, starting with the group with one contact. The only exception here is MP2.5 for the group with $n = 1$, where the errors are below 0.1 kJ/mol for all the considered systems. Therefore, the RPA methods based on HF orbitals are generally more effective than MP methods for describing the trimers with close contacts.

In summary, the RPA methods based on DFT orbitals are less accurate than those based on HF orbitals for the 3-body terms of the considered systems. This arises from the poor performance of the RPA methods evaluated with DFT orbitals when the molecules in trimers are in close contacts. Among the RPA variants, the RPA+MBPT3(HF) method generally demonstrates the best performance for the 3-body terms of the considered systems. Among the WFT methods, MP2.5 is a good choice for describing the 3-body contributions.

6.3 Four-body terms

We now move to discuss the 4-body terms. The 4-body contributions of the energy components in all methods for all the considered systems are relatively small. Moreover, the calculations of the 4-body energy components are computationally demanding. Therefore, we only examine the accuracy of the RPA and RPA+RSE methods based on DFT orbitals, MP2, and CCSD methods for the 4-body terms, shown in Table A.15 in Attachments and Fig. 6.5.

The 4-body CCSD(T) reference energy is around -0.03 kJ/mol for acetylene/II, and -0.10 kJ/mol for other systems (see Table A.15). These small values come from small contributions of the energy components in the CCSD(T) calculations, and they tend to partially cancel each other. Considering first the RPA method, we find that the 4-body RPA energies based on KS/PBE and KS/PBE0 are too positive and yield large errors when compared with the CCSD(T) refer-

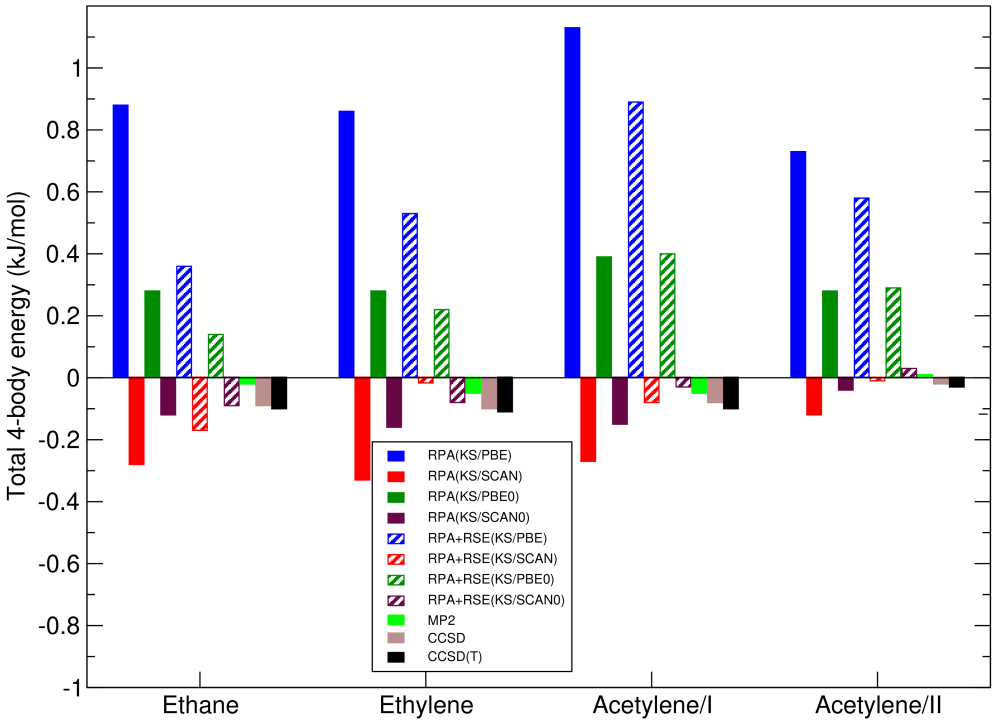

Figure 6.5: The four-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data.

ence data. For example, the error is larger than 1 kJ/mol for acetylene/I when KS/PBE is used. These large errors arise from both mean field and correlation energies. Indeed, the EXX and RPA correlation energies based on KS/PBE and KS/PBE0 are largely positive, and they do not tend to cancel each other for all the considered systems (see Table A.3 and Table A.5). The only exception here is the RPA correlation energy of ethane, which is close to zero (Table A.5). When using KS/SCAN and KS/SCAN0 orbitals, the 4-body RPA energies give smaller errors (see Fig. 6.5). Remarkably, the 4-body RPA(KS/SCAN0) energies are nearly identical to the reference data, with the errors below 0.05 kJ/mol for all considered systems. These small errors mainly originate from the contributions of the EXX component because the 4-body RPA correlation energies based on KS/SCAN and KS/SCAN0 orbitals are almost zero (Table A.5).

The addition of the RSE corrections into RPA generally improves the accuracy of the RPA method. The largest improvement is observed for KS/PBE orbitals, while for other orbitals, the change is less than 0.2 kJ/mol. However, we find that for the two forms of acetylene, RPA+RSE based on KS/PBE0 and KS/SCAN0 orbitals show either similar or even worse results compared to RPA. These changes are very small, and may come from numerical noise of the 4-body RSE contributions. Overall, KS/SCAN and KS/SCAN0 still give better agreement than KS/PBE and KS/PBE0 for RPA+RSE when compared to the CCSD(T) reference.

Among the WFT methods, CCSD shows remarkable accuracy, with the errors of only 0.01 kJ/mol for all the considered systems. The reason for this is that the 4-body (T) energies are very small, and can be safely neglected. The errors in MP2 are larger, but less than 0.10 kJ/mol for all the considered systems. These

errors are comparable to those observed for RPA+RSE when based on KS/SCAN and KS/SCAN0 orbitals.

In summary, KS/SCAN0 orbitals tend to yield superior results compared to other orbitals for both RPA and RPA+RSE methods. We expect that the accuracy of these methods will be improved when using HF orbitals, which should be explored in future studies. The very small contributions of the 4-body terms can make the assessment of the examined methods unreliable. Therefore, further tests on systems with significant 4-body contributions should be carried on to thoroughly assess the accuracy of the examined methods.

6.4 Total binding energies

The accuracy of the approximate methods for describing each n-body contribution was discussed in detail. The question now is how the errors of n-body contributions affect the results of total binding energies. The errors from n-body contributions can accumulate or cancel each other, depending on their magnitudes and signs. We will discuss this in this section by comparing the total RPA and MP2 binding energies to the CCSD(T) reference. The total binding energies are summarized in Table A.16 in Attachments, and their relative deviations from the reference are presented in Fig. 6.6.

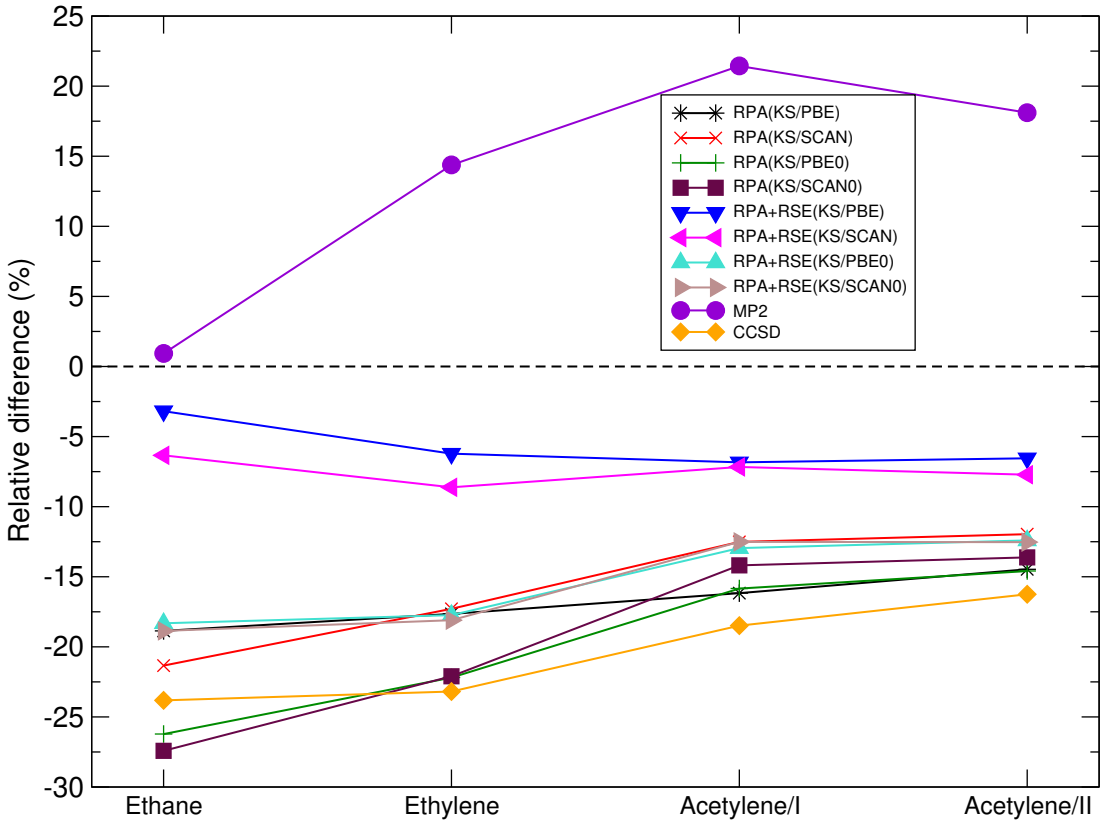

Figure 6.6: Relative difference of the binding energies obtained from the approximate methods with respect to the reference data.

We start to discuss the RPA method. As expected, the binding energies of RPA without RSE corrections underestimate significantly the reference data, with

better agreement observed for the values based on KS/PBE and KS/SCAN orbitals compared to those based on KS/PBE0 and KS/SCAN0 orbitals. The range of errors is between −22% and −12% for KS/PBE and KS/SCAN, and around −27 to −14% for KS/PBE0 and KS/SCAN0. For RPA based on KS/PBE and KS/PBE0 orbitals, the largely positive errors in the 4-body terms are partly cancelled by the negative errors in the 3-body terms, resulting in overall negative errors. Then these negative errors cancel partly the positive errors in the 2-body terms, leading to the total underestimated binding energies. Therefore, the errors in RPA based on KS/PBE and KS/PBE0 depend strongly on error cancellation between n-body terms. However, for RPA based on KS/SCAN and KS/SCAN0, the effect of error cancellation between n-body terms is smaller. The 4-body RPA energies based on these orbitals are close to reference, and have very small errors. Thus, the error cancellation mainly occurs when the negative errors in the 3-body terms counteract the positive errors in the 2-body terms. This finding indicates that RPA based on KS/SCAN and KS/SCAN0 orbitals is better suited for the correction scheme compared to RPA based on KS/PBE and KS/PBE0 orbitals. Interestingly, due to these error cancellations, the relative deviations of the RPA energies based on KS/PBE and KS/SCAN differ by only a few percent, despite significant differences in their n-body terms. We observe the same behavior for RPA based on KS/PBE0 and KS/SCAN0 orbitals. The accuracy of the RPA method is then improved significantly when the RSE corrections are included. The errors in RPA+RSE compared to the CCSD(T) reference are from approximately −9% to −3% for KS/PBE and KS/SCAN orbitals and from −19% to −12% for KS/PBE0 and KS/SCAN0 orbitals.

We now turn to discuss the WFT methods, starting with MP2. We note that the errors in the 4-body MP2 energies are very small for all considered systems (below 0.1 kJ/mol), thus their contributions on the final deviation are minimal. For ethane, MP2 predicts accurately binding energy with a relative error of below 1%. This exceptional accuracy arises from error cancellation between a positive error in the 2-body terms (around 1 kJ/mol) and a negative error in the 3-body terms (around −1.4 kJ/mol). The latter error is partly due to the absence of three-body correlations. For the other systems, we find that MP2 overestimates the binding energies by at least 14%. Although the negative error in the 3-body terms for these systems is somewhat smaller than that for ethane, an inaccurate description with second-order perturbation theory for the systems with π electrons results in significantly negative errors in the 2-body terms. Consequently, the combination of the negative errors in both 2- and 3-body terms leads to a substantially overestimated total MP2 binding energy for these systems.

For all the considered systems, the relative errors obtained from CCSD are generally larger than those observed for RPA based on all orbitals. The only exception here is ethane, where CCSD performs better than RPA based on KS/PBE0 and KS/SCAN0 orbitals. The errors in the 3-body CCSD energies are relatively small (below −0.3 kJ/mol), while the errors in the 4-body terms are negligible. Thus, the notably large errors in CCSD mainly come from the positive errors in the 2-body terms, exceeding 4 kJ/mol for all the considered systems due to the absence of the 2-body (T) component.

Overall, MP2 performs well for ethane but shows significant deterioration for π-electron systems. Conversely, RPA offers a good description for both ethane

and π-electron systems, with RPA+RSE(KS/PBE) being the best choice due to its accuracy. The overall error in the examined methods is strongly influenced by the error cancellation between n-body terms.

6.5 Correction scheme

The calculations of the binding energies for large molecular solids at the CCSD(T) level are currently highly computationally demanding. However, we have shown that the difference in binding energy between the CCSD(T) reference and examined methods occurs mainly at short distances. This motivated us to assess the accuracy of the correction scheme where a portion of n-body contributions, obtained from CCSD(T) data at large distances, is replaced with results obtained from approximate methods. The details of this scheme was discussed in chapter 1. It is expected that the correction scheme performs best when the n-body contributions from the approximate methods are close to the CCSD(T) values. Moreover, the performance of this scheme depends on the error cancellation of the n-body contributions. Therefore, our results show that standard DFT functionals are not suitable for this scheme, and more accurate methods are needed.

The results are relatively similar for all the considered systems, and therefore we focus on the results of ethylene in Table 6.2 as an example, and mention the results of other systems summarized in Attachments (Tables A.17–A.19) when necessary. Let us explain the percentage (%) in Table 6.2 more clearly. This percentage is the fraction of the n-body contributions that are calculated using the approximate method. The rest of the contributions is obtained at the CCSD(T) level. When replacing the contributions, we start from those with the largest distance. Thus, for a percentage of 100% all the CCSD(T) contributions are replaced by the results of the approximate method. A percentage of 10% means that we use 10% of the results from the approximate methods to replace the CCSD(T) results at large distances while retaining 90% of the CCSD(T) results at short distances. To evaluate the performance of the correction scheme, we test how the results differ when the number of contributions obtained at the more approximate level is increased.

Let us start with the 2-body terms. Among the methods tested for this scheme, we expect that 2-body MP2 is a good choice for ethane, while 2-body RPA+RSE(KS/PBE) is a better choice for the other systems. As shown in Table 6.2, substituting 10% of the 2-body CCSD(T) energy with either the 2-body RPA+RSE energy or the 2-body MP2 energy yields negligible errors. The errors are slightly larger when the percentage is increased to 50%, but they remain very small, below 0.03 kJ/mol. Even with a 70% substitution, the errors remain minimal, staying below 0.1 kJ/mol. The errors tend to be smaller when using 2-body MP2 for replacement compared to 2-body RPA+RSE. However, this trend is reversed for the two forms of acetylene, where MP2 predicts wrongly their 2-body binding energies (see Table A.18 and Table A.19 in Attachments). When considering 2-body RPA+RSE as the replacement, we find that the errors increase in the following order of orbitals: KS/PBE < KS/SCAN < KS/PBE0 < KS/SCAN0 < HF. This observation is consistent with the conclusion that KS/PBE orbitals are the best choice for RPA+RSE in the 2-body terms.

Table 6.2: The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for ethylene. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods.

Percentage	n-body	RPA+RSE					MP2
		KS/PBE	KS/SCAN	KS/PBE0	KS/SCAN0	HF	–
100%	2-body	2.34	3.17	4.81	5.23	8.57	−2.13
	3-body	−1.63	−1.22	−1.19	−1.21	−0.44	−1.11
	4-body	0.65	−0.06	0.33	0.04	0.02	0.06
	Total	1.36	1.89	3.95	4.05	8.32	−3.18
90%	2-body	0.11	0.18	0.20	0.23	0.35	−0.10
	3-body	−0.40	−0.34	−0.21	−0.23	0.05	−0.03
	4-body	0.57	−0.06	0.29	0.03	0.05	0.06
	Total	0.28	−0.23	0.28	0.03	0.45	−0.08
70%	2-body	0.02	0.04	0.04	0.05	0.08	−0.03
	3-body	−0.10	−0.10	−0.07	−0.08	−0.04	−0.08
	4-body	0.42	−0.04	0.21	0.02	0.02	0.04
	Total	0.33	−0.10	0.19	−0.01	0.06	−0.07
50%	2-body	0.01	0.02	0.01	0.02	0.03	−0.01
	3-body	−0.05	−0.05	−0.02	−0.03	0.02	0.00
	4-body	0.24	0.00	0.12	0.02	−0.01	0.02
	Total	0.20	−0.03	0.11	0.01	0.05	0.01
30%	2-body	0.00	0.01	0.01	0.01	0.01	−0.01
	3-body	−0.02	−0.02	−0.01	−0.01	−0.01	0.00
	4-body	0.10	0.01	0.05	0.01	0.00	0.00
	Total	0.08	−0.01	0.04	0.00	0.01	0.00
10%	2-body	0.00	0.00	0.00	0.00	0.00	0.00
	3-body	−0.01	0.00	0.00	−0.01	−0.01	0.00
	4-body	0.02	0.00	0.01	0.00	0.01	0.00
	Total	0.02	0.00	0.01	0.00	0.00	−0.01

For the 3-body terms, the errors remain below 0.1 kJ/mol when replacing 70% of the reference data with values from the approximate methods. These errors decrease further, to less than 0.05 kJ/mol, when the substitution fraction is below 50%. The correction scheme works more effectively when using 3-body RPA+RSE based on HF orbitals or 3-body MP2 compared to 3-body RPA+RSE based on the DFT orbitals. This is again consistent with the performance of the approximate methods for the 3-body terms relative to the CCSD(T) reference data.

In the case of the 4-body terms, the errors remain below 0.07 kJ/mol when 90% of the CCSD(T) results is treated with RPA+RSE based on meta-GGA and its hybrids, or HF. Similarly, substituting with the MP2 results yields the same level of errors. However, relatively large errors are observed when using RPA+RSE based on KS/PBE and KS/PBE0, starting from a 30% replacement. This is due to the poor performance of the RPA+RSE method based on KS/PBE and KS/PBE0 for the 4-body terms.

Clearly, there is an error in the correction scheme for each n-body contribution. Fortunately, the errors in different n-body contributions tend to cancel each other. For instance, when 90% of replacement is done by RPA+RSE(KS/SCAN0), the positive error in the 2-body terms can be cancelled completely by the negative error in the 3-body terms. This make the correction scheme more useful when considering the total binding energies. Moreover, we find that the methods which are the best for describing total binding energies may not yield the best precision for the correction scheme. For example, while the RPA+RSE(KS/SCAN0) and MP2 methods describe poorly for the total binding energy of ethylene, they exhibit stronger error cancellation between n-body terms compared to other methods when using the correction scheme, resulting in smaller overall errors. Indeed, the errors in the total binding energy is very small (below 0.1 kJ/mol) when up to 90% of RPA+RSE(KS/SCAN0) or MP2 is used for replacement (see Table 6.2).

Overall, the correction scheme appears to be highly promising, and its accuracy varies depending on the percentage of the replacement and the approximate methods considered. This scheme proves particularly advantageous for molecular solids, where the large number of n-body terms can be efficiently evaluated using simpler methods, resulting in significant reductions in computational requirements.

6.6 Dispersion-corrected DFT methods

The CCSD(T) dataset for all the considered systems is valuable as it can be used to test the accuracy of other methods. DFT has been widely used to study molecular solids. However, it is known that standard DFT functionals can not describe exactly long-range van der Waals interactions, which are important for high-dispersion systems. This issue can be addressed by including dispersion interactions in DFT. The question now is how well these dispersion-corrected DFT methods describe binding energies of molecular solids considered. To this end, we first obtained binding energies of the considered systems using the PBC approach with PBE0-D2 [133], PBE0-D3 [125], PBE0-D3BJ [125, 136], PBE0-rsMBD [147], and PBE0-TS [145] functionals. In these calculations, we used the Coulomb cut-off technique to avoid extrapolations of the energy with k-points and cell volume. Then, a box of $25 \times 25 \times 25$ Å^3 was used for the calculations of all the isolated molecules. For the calculations of solids, a $6 \times 6 \times 6$ k-point set was used for solid ethane and ethylene, and a $4 \times 4 \times 4$ k-point set was used for the two forms of solid acetylene. The basis-set cut-off was set to 1000 eV and the "hard" PAW data sets were used. We note that these settings were found to obtain converged energy values for our considered systems. Other systems may require different settings.

The results of the dispersion-corrected DFT methods compared to the reference CCSD(T) data are shown in Fig. 6.7. It is evident that all selected DFT models tend to overestimate the reference data to a different extent. Among them, the PBE0-D2 model shows the best agreement for all the systems examined, except for ethane, where the PBE0-D3BJ model exhibits the smallest error. Conversely, the PBE0-TS model overestimates strongly the reference data for all the considered systems. Clearly, the performance of the dispersion-corrected DFT

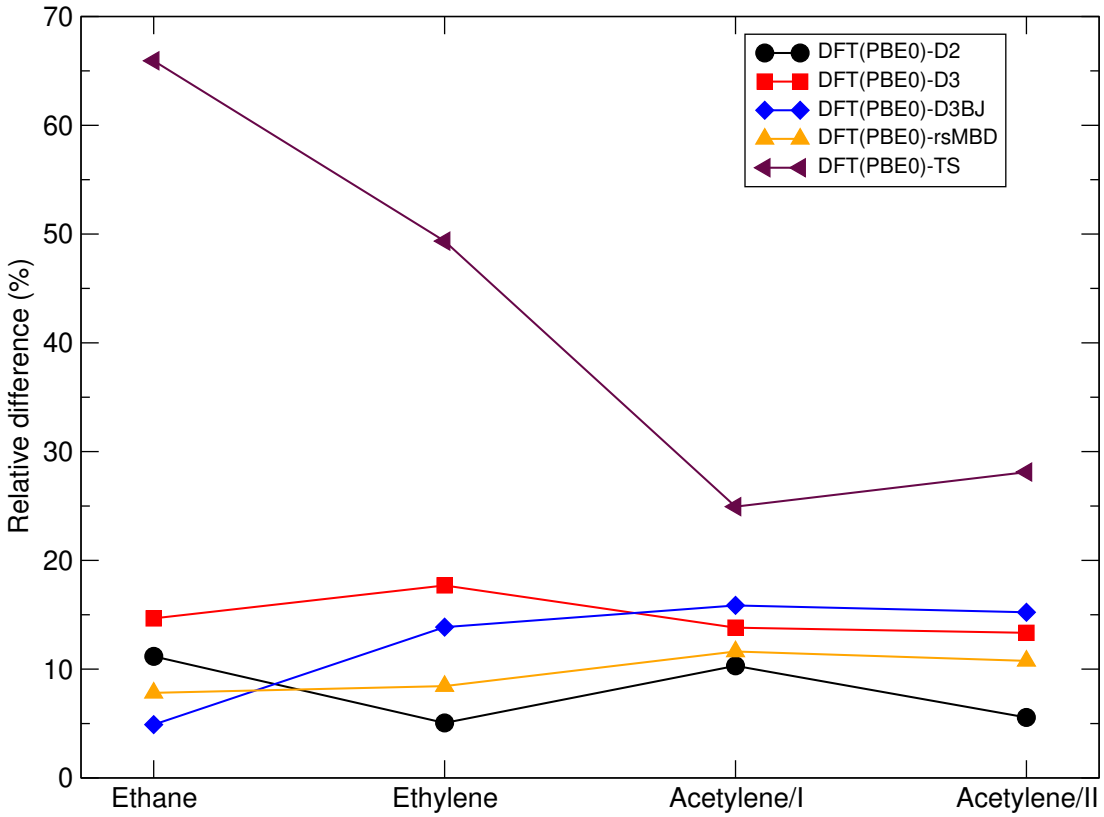

Figure 6.7: Relative difference of the binding energies obtained from several dispersion-corrected DFT methods with respect to the reference data.

methods is influenced by the specific parameters used in dispersion correction.

The performance of the studied DFT methods, except for PBE0-TS, typically surpasses that of the RPA method without RSE corrections. Upon incorporating RSE corrections, the RPA+RSE method using KS/PBE and KS/SCAN states demonstrates superior performance compared to the studied DFT methods. However, both the PBE0-D2 and PBE0-rsMBD methods perform better than the RPA+RSE method based on KS/PBE0 and KS/SCAN0 for all the systems considered.

Overall, dispersion-corrected DFT methods can offer a valuable tool for obtaining reliable binding energies of molecular solids. However, the accuracy of these methods depends strongly on the choice of parameters in the dispersion correction and the characteristics of the considered systems. Therefore, we recommend that a systematic test should be carried to assess thoroughly the reliability of these methods before employing them in actual calculations for binding energies of molecular solids.

Conclusion

In this thesis we developed an efficient setup for obtaining reference binding energies from MBE at the CCSD(T) level of four crystalline hydrocarbons. We then used the reference CCSD(T) results to examine the accuracy of several electronic structure methods for describing n-body contributions and total binding energies of the considered systems. Moreover, we compared the binding energies obtained from MBE and PBC approaches to evaluate the reliability of our MBE approach. The main findings from obtained data can be summarized as follows.

It is challenging to achieve converged values for binding energies of molecular solids with respect to numerical parameters in both MBE and PBC calculations. In MBE, the uncertainties in the energy components arise from the problem with the cut-off distance convergence, and the accuracy of the correlation energies is further affected by the basis set convergence issue. In PBC, using the Coulomb cut-off technique is essential for accelerating the convergence of the energy components with k-points and cell volume, yet the calculations of correlation energies are limited by computational and memory requirements. The values of the energy components obtained with MBE and PBC approaches agree well, and the differences in the energy components between two approaches are consistent with uncertainties estimated for the values. We suggest that combining the mean field energies from PBC with the correlation energies from MBE is a promising approach to achieving total binding energies with a high precision.

Within the MBE approach, the basis set convergence is notably slower for the RPA correlation energies than for the correlation energies based on WFT. The inclusion of the SOSEX corrections alongside RPA can improve the slow basis set convergence of the RPA correlation energies. Further research is needed to explore approaches for addressing this issue effectively. Obtaining converged values of the energy components at short distances requires a large basis set or CBS, whereas smaller basis sets, such as AVTZ or even AVDZ, are sufficient for the energy components at large distances. The contribution of the mean field energies and SOSEX corrections mainly come from short-range interactions, yet the contributions of the correlation energies at large distances need to be considered in the total energy summed over all distances. These findings indicate some strategies which can be used to save computational time.

The orbital inputs influence significantly the results of the energy components in the RPA calculations. The energy changes in all n-body terms are much larger when transitioning from KS/PBE to KS/PBE0 compared to transitioning from KS/SCAN to KS/SCAN0. The orbital inputs also affect the basis set convergence of the RPA correlation energies, with the values based on HF orbitals converging faster than those based on DFT orbitals. Thus, the non-additive RPA(HF) correlation energies can be obtained with sufficiently large basis set without extrapolations to CBS.

The accuracy of the examined methods depends on the n-body terms and the systems considered. None of the methods tested achieves errors below 1 kJ/mol for both 2-body interactions and non-additive contributions. For the RPA and beyond-RPA methods, KS/PBE orbitals demonstrate better performance than other orbitals for the 2-body terms, whereas HF orbitals are significantly supe-

rior to DFT orbitals for the non-additive terms. Specifically, rPT2(KS/PBE) achieves almost a benchmark-level accuracy for the 2-body terms, but inaccurately predicts the 3-body contributions. Conversely, RPA+MBPT3(HF) is the most accurate method for the 3-body terms, but systematically underestimates the 2-body interactions. Therefore, further beyond-RPA methods need to be investigated to find a single method that yields uniform accuracy for both dimers and non-additive contributions. For the WFT methods, MP3 is a good choice for the 2-body terms, while MP2.5 is suitable for the 3-body terms.

The n-body components in the RPA calculations strongly depend on the choice of orbital inputs. However, the errors arising from these n-body contributions tend to cancel each other. Consequently, the total binding energies based on KS/PBE and KS/SCAN, as well as those based on KS/PBE0 and KS/SCAN0, are similar. The MP2 method predicts the total binding energy of ethane almost accurately, but for the other systems, RPA with additional corrections is the best choice among the examined methods.

Overall, we find that the description of binding energies of molecular solids with strong dispersion interactions remains a challenge for approximate electronic structure methods. The many-body-resolved binding energies provide more detailed insights into the sources of errors for a given method than assessment based solely on total binding energies. RPA with additional corrections demonstrates a good choice for describing the binding energies of the considered systems. However, further improvements are required to obtain a RPA variant with uniform accuracy for all n-body contributions. The CCSD(T) data for the solids studied in this work can be used as a benchmark for the development of novel low-scaling approaches. The proposed correction scheme offers substantial savings in computational cost and can be used for further work in molecular solids.

Bibliography

[1] W. T. M. Mooij, B. P. van Eijck, and J. Kroon. Ab initio crystal structure predictions for flexible hydrogen-bonded molecules. *Journal of the American Chemical Society*, 122(14):3500–3505, 2000.

[2] S. L. Price. Predicting crystal structures of organic compounds. *Chemical Society Reviews*, 43(7):2098–2111, 2014.

[3] G. J. O. Beran. Modeling polymorphic molecular crystals with electronic structure theory. *Chemical Review*, 116:5567–5613, 2016.

[4] A. H. Mazurek, Ł. Szeleszczuk, and P. M. Pisklak. Periodic DFT calculations–review of applications in the pharmaceutical sciences. *Pharmaceutics*, 12(5), 2020.

[5] C. Y. Ma, A. A. Moldovan, A. G. P. Maloney, and K. J. Roberts. Exploring the CSD drug subset: An analysis of lattice energies and constituent intermolecular interactions for the crystal structures of pharmaceuticals. *Journal of Pharmaceutical Sciences*, 112(2):435–445, 2023.

[6] A. J. Cruz-Cabeza, S. M. Reutzel-Edens, and J. Bernstein. Facts and fictions about polymorphism. *Chemical Society Reviews*, 44(23):8619–8635, 2015.

[7] S. L. Morissette, O. Almarsson, M. L. Peterson, J. F. Remenar, M. J. Read, A. V. Lemmo, S. Ellis, M. J. Cima, and C. R. Gardner. High-throughput crystallization: Polymorphs, salts, co-crystals and solvates of pharmaceutical solids. *Advanced Drug Delivery Reviews*, 56(3):275–300, 2004.

[8] D. E. Braun, R. M. Bhardwaj, A. J. Florence, D. A. Tocher, and S. L. Price. Complex polymorphic system of gallic acid-five monohydrates, three anhydrates, and over 20 solvates. *Crystal Growth & Design*, 13(1):19–23, 2013.

[9] E. Grothe, H. Meekes, E. Vlieg, J. H. ter Horst, and R. de Gelder. Solvates, salts, and cocrystals: A proposal for a feasible classification system. *Crystal Growth & Design*, 16(6):3237–3243, 2016.

[10] J. Haleblian and W. McCrone. Pharmaceutical applications of polymorphism. *Journal of Pharmaceutical Sciences*, 58(8):911–929, 1969.

[11] D. Braga, L. Casali, and F. Grepioni. The relevance of crystal forms in the pharmaceutical field: Sword of damocles or innovation tools? *International Journal of Molecular Sciences*, 23(16), 2022.

[12] A. Chettri, A. Subba, G. P. Singh, and P. P. Bag. Pharmaceutical co-crystals: A green way to enhance drug stability and solubility for improved therapeutic efficacy. *Journal of Pharmacy and Pharmacology*, 76(1):1–12, 2023.

[13] J. Hoja, H.-Y. Ko, M. A. Neumann, R. Car, R. A. DiStasio Jr., and A. Tkatchenko. Reliable and practical computational description of molecular crystal polymorphs. *Science Advances*, 5(1):eaau3338, 2019.

[14] G. J. O. Beran. Frontiers of molecular crystal structure prediction for pharmaceuticals and functional organic materials. *Chemical Science*, 14:13290–13312, 2023.

[15] S. L. Price. Computer crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. *Accounts of Chemical Research*, 42(1):117–126, 2009.

[16] A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, and A. Michaelides. Fast and accurate quantum Monte Carlo for molecular crystals. *Proceedings of the National Academy of Sciences of the United States of America*, 115(8):1724–1729, 2018.

[17] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G. K. Chan. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science*, 345(6197):640–643, 2014.

[18] J. Nyman and G. M. Day. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm*, 17(28):5154–5165, 2015.

[19] C. Greenwell, J. L. McKinley, P. Zhang, Q. Zeng, G. Sun, B. Li, S. Wen, and G. J. O. Beran. Overcoming the difficulties of predicting conformational polymorph energetics in molecular crystals via correlated wavefunction methods. *Chemical Science*, 11:2200–2214, 2020.

[20] G. J. O. Beran, E. S. Wright, C. Greenwell, and J. A. Cruz-Cabeza. The interplay of intra- and intermolecular errors in modeling conformational polymorphs. *The Journal of Chemical Physics*, 156(10):104112, 2022.

[21] J. Klimeš. Lattice energies of molecular solids from the random phase approximation with singles corrections. *The Journal of Chemical Physics*, 145(9):094506, 2016.

[22] H. K. Buchholz and M. Stein. Accurate lattice energies of organic molecular crystals from periodic turbomole calculations. *Journal of Computational Chemistry*, 39(19):1335–1343, 2018.

[23] J. Hofierka and J. Klimes. Binding energies of molecular solids from fragment and periodic approaches. *Electronic Structure*, 3(3):034010, 2021.

[24] D. Alfè, A. P. Bartók, G. Csányi, and M. J. Gillan. Communication: Energy benchmarking with quantum Monte Carlo for water nano-droplets and bulk liquid water. *The Journal of Chemical Physics*, 138(22):221102, 2013.

[25] G. J. O. Beran. A new era for ab initio molecular crystal lattice energy prediction. *Angewandte Chemie International Edition*, 54(2):396–398, 2015.

[26] D. McDonagh, C.-K. Skylaris, and G. M. Day. Machine-learned fragment-based energies for crystal structure prediction. *Journal of Chemical Theory and Computation*, 15(4):2743–2758, 2019.

[27] C. T. Sargent, D. P. Metcalf, Z. L. Glick, C. H. Borca, and C. D. Sherrill. Benchmarking two-body contributions to crystal lattice energies and a range-dependent assessment of approximate methods. *The Journal of Chemical Physics*, 158(5):054112, 2023.

[28] K. M. Herman and S. S. Xantheas. A formulation of the many-body expansion (MBE) for periodic systems: Application to several ice phases. *The Journal of Physical Chemistry Letters*, 14:989–999, 2023.

[29] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi. Towards an exact description of electronic wavefunctions in real solids. *Nature*, 493(7432):365–370, 2013.

[30] C. Červinka, M. Fulem, and K. Růžička. CCSD(T)/CBS fragment-based calculations of lattice energy of molecular crystals. *The Journal of Chemical Physics*, 144(6):064505, 2016.

[31] H. C. Borca, L. Z. Glick, P. D. Metcalf, A. L. Burns, and D. C. Sherrill. Benchmark coupled-cluster lattice energy of crystalline benzene and assessment of multi-level approximations in the many-body expansion. *The Journal of Chemical Physics*, 158(23):234102, 2023.

[32] S. Kristyán and P. Pulay. Can (semi)local density functional theory account for the London dispersion forces? *Chemical Physics Letters*, 229(3):175–180, 1994.

[33] J. Moellmann and S. Grimme. DFT-D3 study of some molecular crystals. *The Journal of Physical Chemistry C*, 118(14):7615–7621, 2014.

[34] A. V. Terentjev, L. A. Constantin, and J. M. Pitarke. Dispersion-corrected PBEsol exchange-correlation functional. *Phyiscal Review B*, 98(21):214108, 2018.

[35] D. Geatches, I. Rosbottom, R. L. M. Robinson, P. Byrne, P. Hasnip, M. I. J. Probert, D. Jochym, A. Maloney, and K. J. Roberts. Off-the-shelf DFT-DISPersion methods: Are they now "on-trend" for organic molecular crystals? *The Journal of Chemical Physics*, 151(4):044106, 2019.

[36] A. J. A. Price, A. Otero-de-la Roza, and E. R. Johnson. XDM-corrected hybrid DFT with numerical atomic orbitals predicts molecular crystal lattice energies with unprecedented accuracy. *Chemical Science*, 14:1252–1262, 2023.

[37] A. M. Reilly and A. Tkatchenko. Seamless and accurate modeling of organic molecular materials. *The Journal of Physical Chemistry Letters*, 4(6):1028–1033, 2013.

[38] G. A. Dolgonos, O. A. Loboda, and A. D. Boese. Development of embedded and performance of density functional methods for molecular crystals. *The Journal of Physical Chemistry A*, 122(2):708–713, 2018.

[39] D. Lu, Y. Li, D. Rocca, and G. Galli. Ab initio calculation of van der Waals bonded molecular crystals. *Physical Review Letters*, 102:206411, 2009.

[40] M. D. Ben, J. Hütter, and J. VandeVondele. Second-order Møller–Plesset perturbation theory in the condensed phase: An efficient and massively parallel gaussian and plane waves approach. *Journal of Chemical Theory and Computation*, 8(11):4177–4188, 2012.

[41] J. Klimeš, M. Kaltak, E. Maggio, and G. Kresse. Singles correlation energy contributions in solids. *The Journal of Chemical Physics*, 143(10):102816, 2015.

[42] B. D. Nguyen, G. P. Chen, M. M. Agee, A. M. Burow, M. P. Tang, and F. Furche. Divergence of many-body perturbation theory for noncovalent interactions of large molecules. *Journal of Chemical Theory and Computation*, 16(4):2258–2273, 2020.

[43] M. Modrzejewski, S. Yourdkhani, and J. Klimeš. Random phase approximation applied to many-body noncovalent systems. *Journal of Chemical Theory and Computation*, 16(1):427–442, 2020.

[44] M. Modrzejewski, S. Yourdkhani, S. Śmiga, and J. Klimeš. Random-phase approximation in many-body noncovalent systems: Methane in a dodecahedral water cage. *Journal of Chemical Theory and Computation*, 17(2):804–817, 2021.

[45] D. Cieśliński, A. M. Tucholska, and M. Modrzejewski. Post-Kohn–Sham random-phase approximation and correction terms in the expectation-value coupled-cluster formulation. *Journal of Chemical Theory and Computation*, 19(19):6619–6631, 2023.

[46] S. Pisana, M. Lazzeri, C. Casiraghi, K. S. Novoselov, A. K. Geim, A. C. Ferrari, and F. Mauri. Breakdown of the adiabatic Born-Oppenheimer approximation in graphene. *Nature materials*, 6(3):198–201, 2007.

[47] A. Szabó and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Books on Chemistry. Dover Publications, New York, 1996.

[48] F. Jensen. *Introduction to Computational Chemistry*. Third edition. John Wiley & Sons Ltd., England, 2017.

[49] R. J. Bartlett and M. Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291–352, 2007.

[50] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, 157(6):479–483, 1989.

[51] R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga. Non-iterative fifth-order triple and quadruple excitation energy corrections in correlated methods. *Chemical Physics Letters*, 165(6):513–522, 1990.

[52] G. R. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, Oxford, 1989.

[53] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136:B864–B871, 1964.

[54] W. Kohn. Nobel lecture: Electronic structure of matter–wave functions and density functionals. *Review of Modern Physics*, 71:1253–1266, 1999.

[55] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 76(12):6062–6065, 1979.

[56] E. H. Lieb. Density functionals for coulomb systems. *International Journal of Quantum Chemistry*, 24(3):243–277, 1983.

[57] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140:A1133–A1138, 1965.

[58] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[59] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048–5079, 1981.

[60] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244–13249, 1992.

[61] J. Andzelm and E. Wimmer. Density functional Gaussian-type-orbital approach to molecular geometries, vibrations, and reaction energies. *The Journal of Chemical Physics*, 96(2):1280–1303, 1992.

[62] L. He, F. Liu, G. Hautier, M. J. T. Oliveira, M. A. L. Marques, F. D. Vila, J. J. Rehr, G.-M. Rignanese, and A. Zhou. Accuracy of generalized gradient approximation functionals for density-functional perturbation theory calculations. *Physical Review B*, 89:064305, 2014.

[63] J. G. Brandenburg, T. Maas, and S. Grimme. Benchmarking DFT and semiempirical methods on structures and lattice energies for ten ice polymorphs. *The Journal of Chemical Physics*, 142(12):124104, 2015.

[64] C. Lee, H. Chen, and G. Fitzgerald. Structures of the water hexamer using density functional methods. *The Journal of Chemical Physics*, 101(5):4472–4473, 1994.

[65] N. Mardirossian and M. Head-Gordon. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Molecular Physics*, 115(19):2315–2372, 2017.

[66] S. Lehtola and M. A. L. Marques. Meta-local density functionals: A new rung on Jacob's ladder. *Journal of Chemical Theory and Computation*, 17(2):943–948, 2021.

[67] T. J. Giese and D. M. York. Density-functional expansion methods: Evaluation of LDA, GGA, and meta-GGA functionals and different integral approximations. *The Journal of Chemical Physics*, 133(24):244107, 2010.

[68] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, 1996.

[69] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785–789, 1988.

[70] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098–3100, 1988.

[71] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler. Performance of various density-functional approximations for cohesive properties of 64 bulk solids. *New Journal of Physics*, 20(6):063020, 2018.

[72] M. J. Gillan, D. Alfè, and A. Michaelides. Perspective: How good is DFT for water? *The Journal of Chemical Physics*, 144(13):130901, 2016.

[73] B. Vlaisavljevich, J. Huck, Z. Hulvey, K. Lee, J. A. Mason, J. B. Neaton, J. R. Long, C. M. Brown, D. Alfè, A. Michaelides, and B. Smit. Performance of van der Waals corrected functionals for guest adsorption in the M2(dobdc) metal–organic frameworks. *The Journal of Physical Chemistry A*, 121(21):4139–4151, 2017.

[74] V. M. Rayón and I. Cabria. Assessment of density functional approximations for N_2 and CO_2 physisorption on benzene and graphene. *Journal of Computational Chemistry*, 43(21):1403–1419, 2022.

[75] E. R. Johnson, A. D. Becke, C. D. Sherrill, and G. A. DiLabio. Oscillations in meta-generalized-gradient approximation potential energy surfaces for dispersion-bound complexes. *The Journal of Chemical Physics*, 131(3):034111, 2009.

[76] S. E. Wheeler and K. N. Houk. Integration grid errors for meta-GGA-predicted reaction energies: Origin of grid errors for the M06 suite of functionals. *Journal of Chemical Theory and Computation*, 6(2):395–404, 2010.

[77] Y. Zhao, N. E. Schultz, and D. G. Truhlar. Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *The Journal of Chemical Physics*, 123(16):161103, 2005.

[78] Y. Zhao, N. E. Schultz, and D. G. Truhlar. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Theory and Computation*, 2(2):364–382, 2006.

[79] Y. Zhao and D. G. Truhlar. Density functional for spectroscopy: No long-range self-interaction error, good performance for rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *The Journal of Physical Chemistry A*, 110(49):13126–13130, 2006.

[80] Y. Zhao and D. G. Truhlar. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *The Journal of Chemical Physics*, 125(19):194101, 2006.

[81] Y. Zhao and D. G. Truhlar. Density functionals with broad applicability in chemistry. *Accounts of Chemical Research*, 41(2):157–167, 2008.

[82] 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. *Theoretical Chemistry Accounts*, 120(1):215–241, 2008.

[83] Y. Zhao and D. G. Truhlar. Construction of a generalized gradient approximation by restoring the density-gradient expansion and enforcing a tight Lieb–Oxford bound. *The Journal of Chemical Physics*, 128(18):184109, 2008.

[84] Y. Zhao and D. G. Truhlar. Exploring the limit of accuracy of the global hybrid meta density functional for main-group thermochemistry, kinetics, and noncovalent interactions. *Journal of Chemical Theory and Computation*, 4(11):1849–1868, 2008.

[85] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha. Erratum: Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Physical Review Letters*, 82(25):5179–5179, 1999.

[86] 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. *Physical Review Letters*, 91(14):146401, 2003.

[87] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun. Workhorse semilocal density functional for condensed matter physics and quantum chemistry. *Physical Review Letters*, 103(2):026403, 2009.

[88] J. Sun, B. Xiao, and A. Ruzsinszky. Communication: Effect of the orbital-overlap dependence in the meta generalized gradient approximation. *The Journal of Chemical Physics*, 137(5):051101, 2012.

[89] J. Sun, A. Ruzsinszky, and J. P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Physical Review Letters*, 115(3):036402, 2015.

[90] J. Tao and Y. Mo. Accurate semilocal density functional for condensed-matter physics and quantum chemistry. *Physical Review Letters*, 117:073001, 2016.

[91] P. Verma and D. G. Truhlar. HLE16: A local Kohn–Sham gradient approximation with good performance for semiconductor band gaps and molecular excitation energies. *The Journal of Physical Chemistry Letters*, 8(2):380–387, 2017.

[92] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew. Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional. *Nature Chemistry*, 8:831–836, 2016.

[93] E. B. Isaacs and C. Wolverton. Performance of the strongly constrained and appropriately normed density functional for solid-state materials. *Physical Review Materials*, 2:063801, 2018.

[94] L. Zheng, M. Chen, Z. Sun, H.-Y. Ko, B. Santra, P. Dhuvad, and X. Wu. Structural, electronic, and dynamical properties of liquid water by ab initio molecular dynamics based on SCAN functional within the canonical ensemble. *The Journal of Chemical Physics*, 148(16):164505, 2018.

[95] H.-D. Saßnick and C. Cocchi. Electronic structure of cesium-based photocathode materials from density functional theory: Performance of PBE, SCAN, and HSE06 functionals. *Electronic structure*, 3:027001, 2021.

[96] A. P. Bartók and J. R. Yates. Regularized SCAN functional. *The Journal of Chemical Physics*, 150(16):161101, 2019.

[97] D. Mejía-Rodríguez and S. B. Trickey. Comment on "Regularized SCAN functional" [J. Chem. Phys. 150, 161101 (2019)]. *The Journal of Chemical Physics*, 151(20):207101, 2019.

[98] A. P. Bartók and J. R. Yates. Response to "Comment on 'Regularized SCAN functional'" [J. Chem. Phys. 151, 207101 (2019)]. *The Journal of Chemical Physics*, 151(20):207102, 2019.

[99] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun. Accurate and numerically efficient r^2SCAN meta-generalized gradient approximation. *The Journal of Physical Chemistry Letters*, 11(19):8208–8215, 2020.

[100] D. Mejía-Rodríguez and S. B. Trickey. Meta-GGA performance in solids at almost GGA cost. *Physical Review B*, 102:121109, 2020.

[101] R. Kingsbury, A. S. Gupta, C. J. Bartel, J. M. Munro, S. Dwaraknath, M. Horton, and K. A. Persson. Performance comparison of r^2SCAN

and SCAN metaGGA density functionals for solid materials via an automated, high-throughput computational workflow. *Physical Review Materials*, 6:013801, 2022.

[102] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun. Construction of meta-GGA functionals through restoration of exact constraint adherence to regularized SCAN functionals. *The Journal of Chemical Physics*, 156(3):034109, 2022.

[103] M. Kothakonda, A. D. Kaplan, E. B. Isaacs, C. J. Bartel, J. W. Furness, J. Ning, C. Wolverton, J. P. Perdew, and J. Sun. Testing the r^{2}SCAN density functional for the thermodynamic stability of solids with and without a van der Waals correction. *ACS Materials Au*, 3(2):102–111, 2023.

[104] S. Dasgupta, E. Lambros, P. J. Perdew, and F. Paesani. Elevating density functional theory to chemical accuracy for water simulations through a density-corrected many-body formalism. *Nature communications*, 12:6359, 2021.

[105] S. Dasgupta, C. Shahi, P. Bhetwal, J. P. Perdew, and F. Paesani. How good is the density-corrected SCAN functional for neutral and ionic aqueous systems, and what is so right about the Hartree–Fock density? *Journal of Chemical Theory and Computation*, 18(8):4745–4761, 2022.

[106] F. Belleflamme and J. Hutter. Radicals in aqueous solution: Assessment of density-corrected SCAN functional. *Physical Chemistry Chemical Physics*, 25:20817–20836, 2023.

[107] J. L. Bao, L. Gagliardi, and D. G. Truhlar. Self-interaction error in density functional theory: An appraisal. *The Journal of Physical Chemistry Letters*, 9(9):2353–2358, 2018.

[108] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Challenges for density functional theory. *Chemical Reviews*, 112(1):289–320, 2012.

[109] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[110] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, 1999.

[111] K. Hui and J.-D. Chai. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *The Journal of Chemical Physics*, 144(4):044114, 2016.

[112] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of Chemical Physics*, 118(18):8207–8215, 2003.

[113] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]. *The Journal of Chemical Physics*, 124(21):219906, 2006.

[114] J. H. Skone, M. Govoni, and G. Galli. Nonempirical range-separated hybrid functionals for solids and molecules. *Physical Review B*, 93:235106, 2016.

[115] S. Jana and P. Samal. A meta-GGA level screened range-separated hybrid functional by employing short range Hartree-Fock with a long range semilocal functional. *Physical Chemistry Chemical Physics*, 20:8999–9005, 2018.

[116] J. Yang, S. Falletta, and A. Pasquarello. Range-separated hybrid functionals for accurate prediction of band gaps of extended systems. *npj Computational Matererial*, 9(108), 2023.

[117] K. Sharkas, J. Toulouse, and A. Savin. Double-hybrid density-functional theory made rigorous. *The Journal of Chemical Physics*, 134(6):064113, 2011.

[118] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, and J. M. L. Martin. Double-hybrid functionals for thermochemical kinetics. *The Journal of Physical Chemistry A*, 112(1):3–8, 2008.

[119] T. Schäfer, B. Ramberger, and G. Kresse. Quartic scaling MP2 for solids: A highly parallelized algorithm in the plane wave basis. *The Journal of Chemical Physics*, 146(10):104101, 2017.

[120] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, and S. Grimme. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics*, 19:32184–32215, 2017.

[121] J. M. L. Martin and G. Santra. Empirical double-hybrid density functional theory: A 'third way' in between WFT and DFT. *Israel Journal of Chemistry*, 60(8-9):787–804, 2020.

[122] G. Santra, R. Calinsky, and J. M. L. Martin. Benefits of range-separated hybrid and double-hybrid functionals for a large and diverse data set of reaction energies and barrier heights. *The Journal of Physical Chemistry A*, 126(32):5492–5505, 2022.

[123] Y. Zhao and D. G. Truhlar. Density functionals for noncovalent interaction energies of biological importance. *Journal of Chemical Theory and Computation*, 3(1):289–300, 2007.

[124] L. F. Molnar, X. He, B. Wang, and K. M. Merz, Jr. Further analysis and comparative study of intermolecular interactions using dimers from the S22 database. *The Journal of Chemical Physics*, 131(6):065102, 2009.

[125] 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. *The Journal of Chemical Physics*, 132(15):154104, 2010.

[126] S. N. Steinmann and C. Corminboeuf. A system-dependent density-based dispersion correction. *Journal of Chemical Theory and Computation*, 6(7):1990–2001, 2010.

[127] A. Otero-de-la Roza, L. M. LeBlanc, and E. R. Johnson. What is 'many-body' dispersion and should I worry about it? *Physical Chemistry Chemical Physics*, 22:8266–8276, 2020.

[128] K. M. Visscher and D. P. Geerke. Deriving a polarizable force field for biomolecular building blocks with minimal empirical calibration. *The Journal of Physical Chemistry B*, 124(9):1628–1636, 2020.

[129] S. Grimme. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25(12):1463–1473, 2004.

[130] B. M. Axilrod and E. Teller. Interaction of the van der Waals type between three atoms. *The Journal of Chemical Physics*, 11(6):299–300, 1943.

[131] Y. Muto. Force between nonpolar molecules. *Journal of the Mathematical Society of Japan*, 17:629, 1943.

[132] C. A. Morgado, J. P. McNamara, I. H. Hillier, N. A. Burton, and M. A. Vincent. Density functional and semiempirical molecular orbital methods including dispersion corrections for the accurate description of noncovalent interactions involving sulfur-containing molecules. *Journal of Chemical Theory and Computation*, 3(5):1656–1664, 2007.

[133] S. Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27(15):1787–1799, 2006.

[134] H. Schröder, J. Hühnert, and T. Schwabe. Evaluation of DFT-D3 dispersion corrections for various structural benchmark sets. *The Journal of Chemical Physics*, 146(4):044115, 2017.

[135] S. Tsuzuki and T. Uchimaru. Accuracy of intermolecular interaction energies, particularly those of hetero-atom containing molecules obtained by DFT calculations with grimme's D2, D3 and D3BJ dispersion corrections. *Physical Chemistry Chemical Physics*, 22:22508–22519, 2020.

[136] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32:1456–21465, 2011.

[137] S. Ehrlich, J. Moellmann, and S. Grimme. Dispersion-corrected density functional theory for aromatic interactions in complex systems. *Accounts of Chemical Research*, 46(4):916–926, 2013.

[138] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics*, 150(15):154122, 2019.

[139] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics*, 150(15):154122, 2019.

[140] E. Caldeweyher, J.-M. Mewes, S. Ehlert, and S. Grimme. Extension and evaluation of the D4 London-dispersion model for periodic systems. *Physical Chemistry Chemical Physics*, 22(16):8499–8512, 2020.

[141] C. J. Nickerson, K. R. Bryenton, A. J. A. Price, and E. R. Johnson. Comparison of density-functional theory dispersion corrections for the DES15K database. *The Journal of Physical Chemistry A*, 127(41):8712–8722, 2023.

[142] A. D. Becke and E. R. Johnson. A density-functional model of the dispersion interaction. *The Journal of Chemical Physics*, 123(15):154101, 2005.

[143] A. D. Becke and E. R. Johnson. Exchange-hole dipole moment and the dispersion interaction: High-order dispersion coefficients. *The Journal of Chemical Physics*, 124(1):014104, 2006.

[144] E. R. Johnson and A. D. Becke. A unified density-functional treatment of dynamical, nondynamical, and dispersion correlations. II. Thermochemical and kinetic benchmarks. *The Journal of Chemical Physics*, 128(12):124105, 2008.

[145] A. Tkatchenko and M. Scheffler. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Physical Review Letters*, 102(7):073005, 2009.

[146] A. Tkatchenko, R. A. DiStasio, Jr., R. Car, and M. Scheffler. Accurate and efficient method for many-body van der Waals interactions. *Physical Review Letters*, 108(23):236402, 2012.

[147] A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr, and A. Tkatchenko. Long-range correlation energy calculated from coupled atomic response functions. *The Journal of Chemical Physics*, 140(18):18A508, 2014.

[148] M. Kim, W. J. Kim, T. Gould, E. K. Lee, S. Lebègue, and H. Kim. uMBD: A materials-ready dispersion correction that uniformly treats metallic, ionic, and van der Waals bonding. *Journal of the American Chemical Society*, 142(5):2346–2354, 2020.

[149] J. Hermann and A. Tkatchenko. Density functional model for van der Waals interactions: Unifying many-body atomic approaches with nonlocal functionals. *Physical Review Letters*, 124:146401, 2020.

[150] B. Winkler and V. Milman. Accuracy of dispersion-corrected density functional theory calculations of elastic tensors of organic molecular structures. *Crystal Growth & Design*, 20(1):206–213, 2020.

[151] A. J. A. Price, K. R. Bryenton, and E. R. Johnson. Requirements for an accurate dispersion-corrected density functional. *The Journal of Chemical Physics*, 154(23):230902, 2021.

[152] D. O'Connor, I. Bier, Y.-T. Hsieh, and N. Marom. Performance of dispersion-inclusive density functional theory methods for energetic materials. *Journal of Chemical Theory and Computation*, 18(7):4456–4471, 2022.

[153] B. Emrem, R. Kempt, K. Finzel, and T. Heine. London dispersion-corrected density functionals applied to van der Waals stacked layered materials: Validation of structure, energy, and electronic properties. *Advanced Theory and Simulations*, 5(7):2200055, 2022.

[154] J. P. Perdew and K. Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings*, 577(1):1–20, 2001.

[155] J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun, and G. I. Csonka. Some fundamental issues in ground-state density functional theory: A guide for the perplexed. *Journal of Chemical Theory and Computation*, 5(4):902–908, 2009.

[156] G. E. Scuseria, T. M. Henderson, and D. C. Sorensen. The ground state correlation energy of the random phase approximation from a ring coupled cluster doubles approach. *The Journal of Chemical Physics*, 129(23):231101, 2008.

[157] D. C. Langreth and J. P. Perdew. The exchange-correlation energy of a metallic surface. *Solid State Communications*, 17(11):1425–1429, 1975.

[158] O. Gunnarsson and B. I. Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Physical Review B*, 13:4274–4298, 1976.

[159] D. C. Langreth and J. P. Perdew. Exchange-correlation energy of a metallic surface: Wave-vector analysis. *Physical Review B*, 15:2884–2901, 1977.

[160] X. Ren, P. Rinke, C. Joas, and M. Scheffler. Random phase approximation with exchange for an accurate description of crystalline polymorphism. *Journal of Materials Science*, 47(21):7447–7471, 2012.

[161] F. Caruso, D. R. Rohr, M. Hellgren, X. Ren, P. Rinke, A. Rubio, and M. Scheffler. Bond breaking and bond formation: How electron correlation is captured in many-body perturbation theory and density-functional theory. *Physical Review Letters*, 110:146403, 2013.

[162] X. Ren, P. Rinke, C. Joas, and M. Scheffler. Random-phase approximation and its applications in computational chemistry and materials science. *Journal of Materials Science*, 47:7447–7471, 2012.

[163] R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255, 1966.

[164] J. Harl and G. Kresse. Accurate bulk properties from approximate many-body techniques. *Physical Review Letters*, 103:056401, 2009.

[165] Y. Li, D. Lu, H.-V. Nguyen, and G. Galli. Van der Waals interactions in molecular assemblies from first-principles calculations. *The Journal of Physical Chemistry A*, 114(4):1944–1952, 2010.

[166] X. Ren, A. Tkatchenko, P. Rinke, and M. Scheffler. Beyond the random-phase approximation for the electron correlation energy: The importance of single excitations. *Physical Review Letters*, 106(15):153003, 2011.

[167] T. Olsen and K. S. Thygesen. Beyond the random phase approximation: Improved description of short-range correlation by a renormalized adiabatic local density approximation. *Physical Review B*, 88(11):115131, 2013.

[168] T. Olsen and K. S. Thygesen. Accurate ground-state energies of solids and molecules from time-dependent density-functional theory. *Physical Review Letters*, 112:203001, 2014.

[169] N. Colonna, M. Hellgren, and S. de Gironcoli. Correlation energy within exact-exchange adiabatic connection fluctuation-dissipation theory: Systematic development and simple approximations. *Physical Review B*, 90(12):125150, 2014.

[170] B. Mussard, D. Rocca, G. Jansen, and J. G. Ángyán. Dielectric matrix formulation of correlation energies in the random phase approximation: Inclusion of exchange effects. *Journal of Chemical Theory and Computation*, 12(5):2191–2202, 2016.

[171] M. Hellgren, N. Colonna, and S. de Gironcoli. Beyond the random phase approximation with a local exchange vertex. *Physical Review B*, 98(4):045117, 2018.

[172] G. P. Chen, M. M. Agee, and F. Furche. Performance and scope of perturbative corrections to random-phase approximation energies. *Journal of Chemical Theory and Computation*, 14(11):5701–5714, 2018.

[173] F. Hummel, A. Grüneis, G. Kresse, and P. Ziesche. Screened exchange corrections to the random phase approximation from many-body perturbation theory. *Journal of Chemical Theory and Computation*, 15(5):3223–3236, 2019.

[174] A. Görling. Hierarchies of methods towards the exact Kohn-Sham correlation energy based on the adiabatic-connection fluctuation-dissipation theorem. *Physical Review B*, 99(23):235120, 2019.

[175] G. P. Chen, V. L. Voora, M. M. Agee, S. G. Balasubramani, and F. Furche. Random-phase approximation methods. *Annual Review of Physical Chemistry*, 68:421–445, 2017.

[176] P. Bleiziffer, A. Heßelmann, and A. Görling. Efficient self-consistent treatment of electron correlation within the random phase approximation. *The Journal of Chemical Physics*, 139(8):084113, 2013.

[177] M. Hellgren, F. Caruso, D. R. Rohr, X. Ren, A. Rubio, M. Scheffler, and P. Rinke. Static correlation and electron localization in molecular dimers from the self-consistent RPA and *GW* approximation. *Physical Review B*, 91(16):165110, 2015.

[178] D. Graf and C. Ochsenfeld. A range-separated generalized Kohn-Sham method including a long-range nonlocal random phase approximation correlation potential. *The Journal of Chemical Physics*, 153(24):244118, 2020.

[179] S. Riemelmoser, M. Kaltak, and G. Kresse. Optimized effective potentials from the random-phase approximation: Accuracy of the quasiparticle approximation. *The Journal of Chemical Physics*, 154(15):154103, 2021.

[180] J. M. Yu, B. D. Nguyen, J. Tsai, D. J. Hernandez, and F. Furche. Self-consistent random phase approximation methods. *The Journal of Chemical Physics*, 155(4):040902, 2021.

[181] X. Ren, P. Rinke, G. E. Scuseria, and M. Scheffler. Renormalized second-order perturbation theory for the electron correlation energy: Concept, implementation, and benchmarks. *Physical Review B*, 88(3):035120, 2013.

[182] A. Grüneis, M. Marsman, J. Harl, L. Schimka, and G. Kresse. Making the random phase approximation to electronic correlation accurate. *The Journal of Chemical Physics*, 131(15):154115, 2009.

[183] A. Heßelmann. Third-order corrections to random-phase approximation correlation energies. *The Journal of Chemical Physics*, 134(20):204107, 2011.

[184] J. C. Slater. Atomic shielding constants. *Physical Review*, 36:57–64, 1930.

[185] S. F. Boys and A. C. Egerton. Electronic wave functions, I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063):542–554, 1950.

[186] S. F. Boys and M. V. Wilkes. The integral formulae for the variational solution of the molecular many-electron wave equation in terms of gaussian functions with direct electronic correlation. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 258(1294):402–411, 1960.

[187] W. J. Hehre, R. Ditchfield, and J. A. Pople. Self-consistent molecular orbital methods. XII. Further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics*, 56(5):2257–2261, 1972.

[188] J. S. Binkley, J. A. Pople, and W. J. Hehre. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *Journal of the American Chemical Society*, 102(3):939–947, 1980.

[189] T. H. Dunning, Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989.

[190] D. E. Woon and T. H. Dunning, Jr. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *The Journal of Chemical Physics*, 98(2):1358–1371, 1993.

[191] D. E. Woon and T. H. Dunning, Jr. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *The Journal of Chemical Physics*, 103(11):4572–4585, 1995.

[192] S. F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[193] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe. State of the art in counterpoise theory. *Chemical Reviews*, 94(7):1873–1885, 1994.

[194] S. Simon, M. Duran, and J. J. Dannenberg. How does basis set superposition error change the potential surfaces for hydrogen-bonded dimers? *The Journal of Chemical Physics*, 105(24):11024–11031, 1996.

[195] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, 1996.

[196] G. Kresse and J. Hafner. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Physical Review B*, 49(20):14251–14269, 1994.

[197] H. J. Monkhorst and J. D. Pack. Special points for brillouin-zone integrations. *Physical Review B*, 13(12):5188–5192, 1976.

[198] J. McClain, Q. Sun, G. K-L. Chan, and T. C. Berkelbach. Gaussian-based coupled-cluster theory for the ground-state and band structure of solids. *Journal of Chemical Theory and Computation*, 13(3):1209–1218, 2017.

[199] K. Liao, T. Shen, X. Z. Li, A. Alavi, and A. Grüneis. Structural and electronic properties of solid molecular hydrogen from many-electron theories. *Physical Review B*, 103(5):054111, 2021.

[200] T. J. Mach and T. D. Crawford. Computing optical rotation via an N-body approach. *Theoretical Chemistry Accounts*, 133(3):1449, 2014.

[201] R. M. Richard, K. U. Lao, and J. M. Herbert. Achieving the CCSD(T) basis-set limit in sizable molecular clusters: Counterpoise corrections for the many-body expansion. *The Journal of Physical Chemistry Letters*, 4(16):2674–2680, 2013.

[202] R. M. Richard, K. U. Lao, and J. M. Herbert. Understanding the many-body expansion for large systems. I. Precision considerations. *The Journal of Chemical Physics*, 141(1):014108, 2014.

[203] K. U. Lao, K.-Y. Liu, R. M. Richard, and J. M. Herbert. Understanding the many-body expansion for large systems. II. Accuracy considerations. *The Journal of Chemical Physics*, 144(16):164105, 2016.

[204] E. E. Dahlke and D. G. Truhlar. Electrostatically embedded many-body expansion for large systems, with applications to water clusters. *Journal of Chemical Theory and Computation*, 3(1):46–53, 2007.

[205] P. J. Bygrave, N. L. Allan, and F. R. Manby. The embedded many-body expansion for energetics of molecular crystals. *The Journal of Chemical Physics*, 137(16):164102, 2012.

[206] S. Wen and G. J. O. Beran. Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly ab initio force field parametrization. *Journal of Chemical Theory and Computation*, 7(11):3733–3742, 2011.

[207] C. Červinka and G. J. O. Beran. Ab initio prediction of the polymorph phase diagram for crystalline methanol. *Chemical Science*, 9(20):4622–4629, 2018.

[208] U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz. Interaction energies of large clusters from many-body expansion. *The Journal of Chemical Physics*, 135(22):224102, 2011.

[209] J. Liu and J. M. Herbert. Pair–pair approximation to the generalized many-body expansion: An alternative to the four-body expansion for ab initio prediction of protein energetics via molecular fragmentation. *Journal of Chemical Theory and Computation*, 12(2):572–584, 2016.

[210] J. Cui, H. Liu, and K. D. Jordan. Theoretical characterization of the $(H_2O)_{21}$ cluster: Application of an n-body decomposition procedure. *The Journal of Physical Chemistry B*, 110(38):18872–18878, 2006.

[211] C. Müller and D. Usvyat. Incrementally corrected periodic local MP2 calculations: I. the cohesive energy of molecular crystals. *Journal of Chemical Theory and Computation*, 9(12):5590–5598, 2013.

[212] M. A. Neumann. Tailor-made force fields for crystal-structure prediction. *The Journal of Physical Chemistry B*, 112(32):9810–9829, 2008.

[213] J. Nyman, O. S. Pundyke, and G. M. Day. Accurate force fields and methods for modelling organic molecular crystals at finite temperatures. *Physical Chemistry Chemical Physics*, 18:15828–15837, 2016.

[214] R. L. M. Robinson, G. Dawn, M. Chris, M. Rebecca, A. G. P. Maloney, K. J. Roberts, M. Alexandru, C. Ernest, P. Klimentina, and D. R. M. Vatvani. Evaluation of force-field calculations of lattice energies on a large

public dataset, assessment of pharmaceutical relevance, and comparison to density functional theory. *Journal of Chemical Information and Modeling*, 59(11):4778–4792, 2019.

[215] M. Thürlemann and S. Riniker. Hybrid classical/machine-learning force fields for the accurate description of molecular condensed-phase systems. *Chemical Science*, 14:12661–12675, 2023.

[216] A. S. Christensen, T. Kubař, Q. Cui, and M. Elstner. Semiempirical quantum mechanical methods for noncovalent interactions for chemical and biochemical applications. *Chemical Reviews*, 116(9):5301–5337, 2016.

[217] J. G. Brandenburg and S. Grimme. Accurate modeling of organic molecular crystals by dispersion-corrected density functional tight binding (DFTB). *The Journal of Physical Chemistry Letters*, 5(11):1785–1789, 2014.

[218] M. Mortazavi, J. G. Brandenburg, R. J. Maurer, and A. Tkatchenko. Structure and stability of molecular crystals with many-body dispersion-inclusive density functional tight binding. *The Journal of Physical Chemistry Letters*, 9(2):399–405, 2018.

[219] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, and S. Grimme. Extended tight-binding quantum chemistry methods. *WIREs Computational Molecular Science*, 11(2):e1493, 2021.

[220] M. Gaus, A. Goez, and M. Elstner. Parametrization and benchmark of DFTB3 for organic molecules. *Journal of Chemical Theory and Computation*, 9(1):338–354, 2013.

[221] A. V. Akimov and O. V. Prezhdo. Large-scale computations in chemistry: A bird's eye view of a vibrant field. *Chemical Reviews*, 115(12):5797–5890, 2015.

[222] B. Santra, J. Klimeš, A. Tkatchenko, D. Alfè, B. Slater, A. Michaelides, R. Car, and M. Scheffler. On the accuracy of van der Waals inclusive density-functional theory exchange-correlation functionals for ice at ambient and high pressures. *The Journal of Chemical Physics*, 139(15):154702, 2013.

[223] A. Otero-de-la Roza and E. R. Johnson. A benchmark for non-covalent interactions in solids. *The Journal of Chemical Physics*, 137(5):054103, 2012.

[224] A. Otero-de-la Roza, L. M. LeBlanc, and E. R. Johnson. Dispersion XDM with hybrid functionals: Delocalization error and halogen bonding in molecular crystals. *Journal of Chemical Theory and Computation*, 15(9):4933–4944, 2019.

[225] T. Bučko, S. Lebègue, T. Gould, and J. G. Ángyán. Many-body dispersion corrections for periodic systems: An efficient reciprocal space implementation. *Journal of Physics: Condensed Matter*, 28(4):045201, 2016.

[226] B. Chen and X. Xu. XO-PBC: An accurate and efficient method for molecular crystals. *Journal of Chemical Theory and Computation*, 16(7):4271–4285, 2020.

[227] F. Stein, J. Hutter, and V. V. Rybkin. Double-hybrid DFT functionals for the condensed phase: Gaussian and plane waves implementation and evaluation. *Molecules*, 25(21):5174, 2020.

[228] K. Sharkas, J. Toulouse, L. Maschio, and B. Civalleri. Double-hybrid density-functional theory applied to molecular crystals. *The Journal of Chemical Physics*, 141(4):044105, 2014.

[229] G. Sansone, B. Civalleri, D. Usvyat, J. Toulouse, K. Sharkas, and L. Maschio. Range-separated double-hybrid density-functional theory applied to periodic systems. *The Journal of Chemical Physics*, 143(10):102811, 2015.

[230] L. Maschio, D. Usvyat, M. Schütz, and B. Civalleri. Periodic local Møller–Plesset second order perturbation theory method applied to molecular crystals: Study of solid NH_3 and CO_2 using extended basis sets. *The Journal of Chemical Physics*, 132(13):134706, 2010.

[231] M. Del Ben, J. VandeVondele, and B. Slater. Periodic MP2, RPA, and boundary condition assessment of hydrogen ordering in ice XV. *The Journal of Physical Chemistry Letters*, 5(23):4122–4128, 2014.

[232] M. Cutini, B. Civalleri, M. Corno, R. Orlando, J. G. Brandenburg, L. Maschio, and P. Ugliengo. Assessment of different quantum mechanical methods for the prediction of structure and cohesive energy of molecular crystals. *Journal of Chemical Theory and Computation*, 12(7):3340–3352, 2016.

[233] S. P. Thomas, P. R. Spackman, D. Jayatilaka, and M. A. Spackman. Accurate lattice energies for molecular crystals from experimental crystal structures. *Journal of Chemical Theory and Computation*, 14(3):1614–1623, 2018.

[234] Y. H. Liang, H.-Z. Ye, and T. C. Berkelbach. Can spin-component scaled MP2 achieve kJ/mol accuracy for cohesive energies of molecular crystals? *The Journal of Physical Chemistry Letters*, 14(46):10435–10441, 2023.

[235] G. J. O. Beran and K. Nanda. Predicting organic crystal lattice energies with chemical accuracy. *The Journal of Physical Chemistry Letters*, 1(24):3480–3487, 2010.

[236] G. Chałasiński, M. M. Szczęśniak, and R. A. Kendall. Supermolecular approach to many-body dispersion interactions in weak van der Waals complexes: He, Ne, and Ar trimers. *The Journal of Chemical Physics*, 101(10):8860, 1994.

[237] S. J. Bintrim, T. C. Berkelbach, and H.-Z. Ye. Integral-direct Hartree-Fock and Møller-Plesset perturbation theory for periodic systems with density fitting: Application to the benzene crystal. *Journal of Chemical Theory and Computation*, 18(9):5374–5381, 2022.

[238] T. Olsen and K. S. Thygesen. Random phase approximation applied to solids, molecules, and graphene-metal interfaces: From van der Waals to covalent bonding. *Physical Review B*, 87:075111, 2013.

[239] T. Olsen, C. E. Patrick, J. E. Bates, A. Ruzsinszky, and K. S. Thygesen. Random phase approximation applied to solids, molecules, and graphene-metal interfaces: From van der Waals to covalent bonding. *npj Computational Materials*, 5, 2019.

[240] H. Eshuis and F. Furche. Basis set convergence of molecular correlation energy differences within the random phase approximation. *The Journal of Chemical Physics*, 136(8):084105, 2012.

[241] G. J. H. van Nes and A. Vos. Single-crystal structures and electron density distributions of ethane, ethylene and acetylene. I. Single-crystal X-ray structure determinations of two modifications of ethane. *Acta Crystallographica Section B*, 34:1947–1956, 1978.

[242] G. J. H. van Nes and A. Vos. Single-crystal structures and electron density distributions of ethane, ethylene and acetylene. III. Single-crystal X-ray structure determination of ethylene at 85 K. *Acta Crystallographica Section B*, 35(11):2593–2601, 1979.

[243] R. K. McMullan, Å. Kvick, and P. Popelier. Structures of cubic and orthorhombic phases of acetylene by single-crystal neutron diffraction. *Acta Crystallographica Section B*, 48:726–731, 1992.

[244] C. R. Groom, I. J. Bruno, M. P. Lighfoot, and S. C. Ward. The cambridge structural database. *Acta Crystallographica B*, 72:171–179, 2016.

[245] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der Waals density functional for general geometries. *Physical Review Letters*, 92(24):246401, 2004.

[246] J. Klimeš, D. R. Bowler, and A. Michaelides. Van der Waals density functionals applied to solids. *Physical Review B*, 83(19):195131, 2011.

[247] G. Román-Pérez and J. M. Soler. Efficient implementation of a van der Waals density functional: Application to double-wall carbon nanotubes. *Physical Review Letters*, 103(9):096102, 2009.

[248] J. Klimeš, D. R. Bowler, and A. Michaelides. Chemical accuracy for the van der Waals density functional. *Journal of Physics: Condensed Matter*, 22(2):022201, 2009.

[249] R. J. Bartlett, I. Grabowski, S. Hirata, and S. Ivanov. The exchange-correlation potential in ab initio density functional theory. *The Journal of Chemical Physics*, 122(3):034104, 2005.

[250] R. J Bartlett. Ab initio DFT and its role in electronic structure theory. *Molecular Physics*, 108(21-23):3299–3311, 2010.

[251] P. Verma and R. J Bartlett. Increasing the applicability of density functional theory. II. Correlation potentials from the random phase approximation and beyond. *The Journal of Chemical Physics*, 136(4):044105, 2012.

[252] C. H. Borca, B. W. Bakr, L. A. Burns, and C. D. Sherrill. Crystalatte: Automated computation of lattice energies of organic crystals exploiting the many-body expansion to achieve dual-level parallelism. *The Journal of Chemical Physics*, 151(14):144103, 2019.

[253] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz. Molpro: A general-purpose quantum chemistry program package. *WIREs Computational Molecular Science*, 2(2):242–253, 2012.

[254] 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. *The Journal of Chemical Physics*, 96(9):6796–6806, 1992.

[255] A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, and P. R. Taylor. Basis set convergence of the interaction energy of hydrogen-bonded complexes. *The Journal of Chemical Physics*, 111(20):9157–9167, 1999.

[256] T. B. Adler, G. Knizia, and H.-J. Werner. A simple and efficient CCSD(T)-F12 approximation. *The Journal of Chemical Physics*, 127(22):221106, 2007.

[257] J. Noga and J. Šimunek. On the one-particle basis set relaxation in R12 based theories. *Chemical Physics*, 356(1-3):1–6, 2009.

[258] W. Kutzelnigg and W. Klopper. Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory. *The Journal of Chemical Physics*, 94(3):1985–2001, 1991.

[259] H.-J. Werner, T. B. Adler, and F. R. Manby. General orbital invariant MP2-F12 theory. *The Journal of Chemical Physics*, 126(16):164102, 2007.

[260] B. Brauer, M. K. Kesharwani, S. Kozuch, and J. M. L. Martin. The S66x8 benchmark for noncovalent interactions revisited: Explicitly correlated ab initio methods and density functional theory. *Physical Chemistry Chemical Physics*, 18(31):20905–20925, 2016.

[261] G. Knizia, T. B. Adler, and H.–J. Werner. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *The Journal of Chemical Physics*, 130(5):054104, 2009.

[262] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558–561, 1993.

[263] P. E. Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953–17979, 1994.

[264] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59(3):1758–1775, 1999.

[265] M. Kaltak, J. Klimeš, and G. Kresse. Low scaling algorithms for the random phase approximation: Imaginary time and laplace transformations. *Journal of Chemical Theory and Computation*, 10(6):2498–2507, 2014.

[266] M. Kaltak, J. Klimeš, and G. Kresse. A cubic scaling algorithm for the random phase approximation: Defect calculations for large Si model structures. *Physical Review B*, 90(5):054115, 2014.

[267] M. Marsman, A. Grüneis, J. Paier, and G. Kresse. Second-order Møller–Plesset perturbation theory applied to extended systems. I. Within the projector-augmented-wave formalism using a plane wave basis set. *The Journal of Chemical Physics*, 130(18):184103, 2009.

[268] J. Klimeš, M. Kaltak, and G. Kresse. Predictive GW calculations using plane waves and pseudopotentials. *Physical Review B*, 90:075125, 2014.

[269] S. Riemelmoser, M. Kaltak, and G. Kresse. Plane wave basis set correction methods for RPA correlation energies. *The Journal of Chemical Physics*, 152(13):134103, 2020.

[270] I. R. Dagg, A. Anderson, W. Smith, M. Missio, C. G. Joslin, and L. A. A. Read. The quadrupole moment of acetylene from collision-induced absorption in a gaseous mixture with argon. *Canadian Journal of Physics*, 66(5):453–459, 1988.

[271] R. Lindh and B. Liu. Accurate ab initio calculations of the quadrupole moment of acetylene. A combined study of basis set, correlation, and vibrational effects. *The Journal of Chemical Physics*, 94(6):4356–4368, 1991.

[272] F. Gygi and A. Baldereschi. Self-consistent Hartree-Fock and screened-exchange calculations in solids - application to silicon. *Physical Review B*, 34(6):4405, 1986.

[273] J. Paier, R. Hirschl, M. Marsman, and G. Kresse. The Perdew-Burke-Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set. *The Journal of Chemical Physics*, 122(23):234102, 2005.

[274] C. A. Rozzi, D. Varsano, A. Marini, E. K. U. Gross, and A. Rubio. Exact coulomb cutoff technique for supercell calculations. *Physical Review B*, 73(20):205119, 2006.

[275] J. Harl and G. Kresse. Cohesive energy curves for noble gas solids calculated by adiabatic connection fluctuation-dissipation theory. *Physical Review B*, 77(4):045136, 2008.

[276] A. Gulans. Towards numerically accurate many-body perturbation theory: Short-range correlation effects. *The Journal of Chemical Physics*, 141(16):164127, 2014.

[277] Z.-H. Yang, H. Peng, J. Sun, and J. P. Perdew. More realistic band gaps from meta-generalized gradient approximations: Only in a generalized Kohn-Sham scheme. *Physical Review B*, 93(20):205205, 2016.

[278] K. Liao and A. Grüneis. Communication: Finite size correction in periodic coupled cluster theory calculations of solids. *The Journal of Chemical Physics*, 145(14):141102, 2016.

[279] J. Paier, X. Ren, P. Rinke, G. E. Scuseria, A. Grüneis, G. Kresse, and M. Scheffler. Assessment of correlation energies based on the random-phase approximation. *New Journal of Physics*, 14:043002, 2012.

[280] M. Macher, J. Klimeš, C. Franchini, and G. Kresse. The random phase approximation applied to ice. *The Journal of Chemical Physics*, 140(8):084502, 2014.

[281] A. Förster. Assessment of the second-order statically screened exchange correction to the random phase approximation for correlation energies. *Journal of Chemical Theory and Computation*, 18(10):5948–5965, 2022.

[282] M. Hellgren and L. Baguet. Random phase approximation with exchange for an accurate description of crystalline polymorphism. *Physical Review Research*, 3(3):033263, 2021.

[283] F.-F. Wang, G. Jenness, W. A. Al-Saidi, and K. D. Jordan. Assessment of the performance of common density functional methods for describing the interaction energies of $(H_2O)_6$ clusters. *The Journal of Chemical Physics*, 132(13):134303, 2010.

[284] M. Hapka, L. Rajchel, M. Modrzejewski, R. Schäffer, G. Chałasiński, and M. M. Szczęśniak. The nature of three-body interactions in DFT: Exchange and polarization effects. *The Journal of Chemical Physics*, 147(8):084106, 2017.

[285] J. E. Z. D. Mirón and M. Stein. A benchmark for non-covalent interactions in organometallic crystals. *Physical Chemistry Chemical Physics*, 24:29338–29349, 2022.

[286] P. Hobza, H. L. Selzle, and E. W. Schlag. Potential energy surface for the benzene dimer. Results of ab initio CCSD(T) calculations show two nearly isoenergetic structures: T-shaped and parallel-displaced. *The Journal of Physical Chemistry*, 100(48):18790–18794, 1996.

[287] S. Grimme. Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *The Journal of Chemical Physics*, 118(20):9095–9102, 2003.

[288] J. Řezáč, Y. Huang, P. Hobza, and G. J. O. Beran. Benchmark calculations of three-body intermolecular interactions and the performance of low-cost electronic structure methods. *Journal of Chemical Theory and Computation*, 11(7):3065–3079, 2015.

List of Figures

3.1 Cut-off distance dependence of the energy differences between the AVQZ and AVTZ basis sets obtained from the 2-body calculations for ethylene. . . . 42
3.2 Cut-off distance dependence of the energy differences between the AVQZ and AVTZ basis sets obtained from the 3-body calculations for ethylene. . . . 43
3.3 Cut-off distance convergence of the 2-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used. . . . 50
3.4 Cut-off distance convergence of the 3-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used. . . . 51
3.5 Cut-off distance convergence of the 4-body DFT (a), EXX (b), RSE (c), and RPA correlation (d) energies. Here data based on KS/PBE were used. . . . 52
3.6 Method for estimating the uncertainty of the 3-body EXX contribution based on KS/PBE of acetylene/II. . . . 53
3.7 Convergence of the EXX(KS/PBE) energy for ethylene solid and isolated molecule as a function of the supercell volume without and with the use of the real-space Coulomb cut-off. For the solid, the supercell volume is obtained as the unit cell volume times the number of k-points and data between 2^3 and 7^3 k-points are shown. For the molecule, the supercell is the simulation cell. A plane-wave basis-set cut-off of 1100 eV was used. . . . 56
3.8 The convergence of the EXX energy for solid ethylene (left) and solid acetylene/II (right) with respect to the number of k-points. The calculations used a plane-wave basis-set cut-off of 1100 eV and the Coulomb cut-off method. Note the different scales on the y-axis. The EXX energy gives the values printed by VASP. . . . 57
3.9 The convergence of the RPA correlation energy of ethylene solid and molecule with k-points and cell volume. RPA was evaluated on KS/PBE (a, b) and KS/PBE0 (c, d) states. A cut-off of 600 eV was used. . . . 59
3.10 The convergence of the binding energy of EXX(KS/PBE) with the plane wave basis sets for ethylene. A k-points with $N_k = 2$ and a cell volume with $a = 8$ were used for the calculations. . . . 61
3.11 The convergence of the binding energy of RPA(KS/PBE0) with the plane wave basis sets for ethane. . . . 61

4.1 The differences in the mean field energies between MBC and PBC approaches. . . . 65
4.2 The differences in the correlation energies between MBC and PBC approaches. . . . 66

6.1 The two-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data. . 74
6.2 The difference between the rPT2(KS/PBE) and CCSD(T) binding energies over distance for the 2-body contributions of ethylene. . . 75
6.3 The three-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data. . 78
6.4 Signed errors in the three-body nonadditive energies of all the considered systems. The results are grouped according to the number of close contacts between the molecules in a trimer. Note the different scales on y axis. 80
6.5 The four-body contributions to the total binding energies of the approximate methods compared to the CCSD(T) reference data. . 82
6.6 Relative difference of the binding energies obtained from the approximate methods with respect to the reference data. 83
6.7 Relative difference of the binding energies obtained from several dispersion-corrected DFT methods with respect to the reference data. 88

List of Tables

2.1 The unit cell volume at experimental equilibrium (V_0), the lattice constants (a, b, c), the number of molecules in the unit cell (Z), and the CSD code of the selected molecular solids. 36

3.1 The basis set convergence of the energy components obtained from the 2-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets. The energy components (E^{DFT}, E^{EXX}, and E^{RSE}) do not depend on the transformation from KS/PBE orbitals to semi/PBE orbitals, and are thus identical. 40

3.2 The basis set convergence of the energy components obtained from the 3-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets. The energy components (E^{DFT}, E^{EXX}, and E^{RSE}) do not depend on the transformation from KS/PBE orbitals to semi/PBE orbitals, and are thus identical. 44

3.3 The basis set convergence of the energy components obtained from the 4-body calculations in different approaches (kJ/mol). The values show differences between the values obtained with the AVQZ and AVTZ basis sets. 45

3.4 The basis sets used for obtaining converged values of the energy components for all the considered systems. 46

3.5 Basis set convergence of the 2-body MP2 correlation energies in kJ/mol for ethylene. Here the energies are given to three decimal digits to be able to show small changes between them. 47

3.6 Cut-off distance (r_{cut}, in Å) and corresponding number (N) of symmetry inequivalent dimers, trimers, and tetramers within the selected cut-off distance for the MBE calculations. 48

3.7 The values of the energy components (kJ/mol) obtained from the 2-body calculations for distant dimers ($10 < r < r_{cut}$) in different approaches. 49

3.8 Cut-off distance interval (R, in Å) and corresponding number (N) of dimers and trimers selected to estimate the uncertainties of energy components. 54

3.9 The estimated uncertainties (in kJ/mol) of the energy components in the first group of the methods for all the considered systems. 54

3.10 The estimated uncertainties (in kJ/mol) of the energy components in the second group of the methods for all the considered systems. 54
3.11 The k-point set and the cell side a in Å used for the PBC calculations. DFT, RPA are settings for KS/PBE and KS/SCAN calculations, DFTh, RPAh are settings for KS/PBE0 and KS/SCAN0 calculations, $\rightarrow$ denotes extrapolation of the energies with steps of 1 k-point grid for solids or 1 Å for isolated molecules. 58
3.12 The values of the binding energies of ethylene (in kJ/mol) obtained from standard and hard potentials. 62

5.1 The mean field energies of ethylene in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE. 69
5.2 The correlation energies of ethylene in kJ/mol obtained from the RPA calculations using the MBE approach. 71

6.1 Total 2-body energies of the examined methods for distant dimers ($10 < r < r_{\text{cut}}$) (kJ/mol) for all the considered systems. 77
6.2 The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for ethylene. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods. 86

A.1 The DFT energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 122
A.2 The HF energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 122
A.3 The EXX energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 123
A.4 The RSE energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 124
A.5 The RPA correlation energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 125
A.6 The MP2 correlation energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems. 125

A.7 The mean field energies of ethane in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE 126
A.8 The mean field energies of acetylene/I in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE 126
A.9 The mean field energies of acetylene/II in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE 127
A.10 The correlation energies of ethane in kJ/mol obtained from the RPA calculations using the MBE approach. 127
A.11 The correlation energies of acetylene/I in kJ/mol obtained from the RPA calculations using the MBE approach. 128
A.12 The correlation energies of acetylene/II in kJ/mol obtained from the RPA calculations using the MBE approach. 128
A.13 Total 2-body energies of the examined methods summed over all distances (kJ/mol) for all the considered systems. 129
A.14 Total 3-body energies of the examined methods summed over all distances (kJ/mol) for all the considered systems. 130
A.15 Total 4-body energies of the examined methods summed over all distances (kJ/mol). 131
A.16 Total n-body energies of the examined methods summed over all distances (kJ/mol). 131
A.17 The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for ethane. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods. 132
A.18 The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for acetylene/I. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods. 133

A.19 The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for acetylene/II. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods. 134

List of Abbreviations

PBC: Periodic boundary conditions
MBE: Many body expansion
BO: Born-Oppenheimer
PES: Potential energy surface
SD: Slater determinant
SCF: Self-consistent field
BSSE: Basis set superposition error
CBS: Complete basis set
BZ: Brillouin zone
DFT: Density functional theory
RDM: Reduced density matrix
UEG: Uniform electron gas
Kohn-Sham: KS
XC: Exchange correlation
LDA: Local density approximation
GGA: Generalized gradient approximation
MGGA: Meta-generalized gradient approximation
DH: Double hybrid
PAW: Projector-augmented-wave
HF: Hatree-Fock
CI: Configuration interaction
CC: Couple cluster
MBPT: Many body perturbation theory
MP: Møller-Plesset
MP2: Second-order Møller-Plesset perturbation theory
MP3: Third-order Møller-Plesset perturbation theory
MP4: Fourth-order Møller-Plesset perturbation theory
RPA: Random phase approximation
ACFD: Adiabatic-connection fluctuation-dissipation
AC: Adiabatic connection
SOSEX: Second-order screened exchange
VASP: Vienna ab-initio simulations package
ML-FF: Machine learning force field
rPT2: Renormalized second-order perturbation theory

List of Publications

P1. K. N. Pham, M. Modrzejewski, and J. Klimeš, "Assessment of random phase approximation and second-order Møller-Plesset perturbation theory for many-body interactions in solid ethane, ethylene, and acetylene," The Journal of Chemical Physics, 158(14), 144119 (2023).

P2. K. N. Pham, M. Modrzejewski, and J. Klimeš, "Contributions beyond direct random-phase approximation in the binding energy of solid ethane, ethylene, and acetylene," The Journal of Chemical Physics, 160(22), 224101 (2024).

P3. K. N. Pham, M. Modrzejewski, and J. Klimeš, in preparation (2024).

A. Attachments

A.1 Chapter 4

The results of the energy components in the RPA and MP2 calculations obtained from both MBE and PBC approaches.

Table A.1: The DFT energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Input	Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	2-body	−5.55	−8.57	−15.88	−15.64
	3-body	6.86	6.86	6.28	5.42
	4-body	−1.81	−1.86	−2.07	−1.18
	Total MBE	−0.5	−3.57	−11.67	−11.40
	PBC	−0.43	−3.52	−11.51	−11.46
KS/SCAN	2-body	−8.31	−11.97	−21.08	−19.17
	3-body	−5.15	−3.08	−1.53	−0.82
	4-body	2.28	1.48	0.57	0.29
	Total MBE	−11.18	−13.57	−22.04	−19.70
	PBC	−11.54	−13.92	−22.33	−19.91
KS/PBE0	2-body	−3.10	−6.52	−14.57	−14.54
	3-body	3.93	4.10	3.67	3.35
	4-body	−1.13	−1.19	−1.30	−0.73
	Total MBE	−0.30	−3.61	−12.38	−11.92
	PBC	−0.25	−3.61	−12.30	−11.98
KS/SCAN0	2-body	−7.58	−11.18	−20.44	−18.79
	3-body	−3.44	−1.92	−0.92	−0.36
	4-body	1.41	0.86	0.29	0.14
	Total MBE	−9.61	−12.24	−21.07	−19.01
	PBC	−9.68	−12.46	−21.29	−19.20

Table A.2: The HF energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
2-body	12.95	9.38	1.64	−2.24
3-body	−0.62	−0.38	−0.90	−0.15
4-body	0.03	0.02	0.03	0.05
Total MBE	12.36	9.02	0.77	−2.34
PBC	12.42	8.96	0.70	−2.37

Table A.3: The EXX energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Input	Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	2-body	19.45	14.22	6.53	1.63
	3-body	−2.95	−2.15	−2.47	−1.36
	4-body	0.89	0.62	0.60	0.41
	Total MBE	17.39	12.69	4.66	0.68
	PBC	17.38	12.63	4.46	0.52
KS/SCAN	2-body	18.02	12.08	3.81	−0.69
	3-body	−0.10	0.46	−0.15	0.54
	4-body	−0.23	−0.31	−0.26	−0.13
	Total MBE	17.69	12.23	3.40	−0.28
	PBC	17.52	12.12	3.34	−0.37
KS/PBE0	2-body	16.03	11.25	3.37	0.95
	3-body	−1.29	−0.79	−1.27	−0.41
	4-body	0.32	0.19	0.15	0.13
	Total MBE	15.06	10.65	2.25	−1.23
	PBC	15.07	10.57	2.10	−1.35
KS/SCAN0	2-body	15.76	10.52	2.32	−1.84
	3-body	−0.14	0.33	−0.22	0.42
	4-body	−0.07	−0.15	−0.14	−0.05
	Total MBE	15.55	10.70	1.96	−1.47
	PBC	15.35	10.57	1.88	−1.53

Table A.4: The RSE energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Input	Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	2-body	−4.46	−3.21	−3.06	−2.40
	3-body	1.46	1.02	0.81	0.60
	4-body	−0.52	−0.33	−0.24	−0.15
	Total MBE	−3.52	−2.52	−2.49	−1.95
	PBC	−3.66	−2.53	−2.46	−1.83
KS/SCAN	2-body	−3.13	−1.47	−0.95	−0.58
	3-body	−0.36	−0.60	−0.67	−0.58
	4-body	0.11	0.16	0.19	0.11
	Total MBE	−3.38	−1.91	−1.43	−1.05
	PBC	−3.47	−1.88	−1.35	−0.90
KS/PBE0	2-body	−1.93	−1.02	−0.79	−0.55
	3-body	0.36	0.15	0.04	0.02
	4-body	−0.14	−0.06	0.01	0.01
	Total MBE	−1.71	−0.93	−0.74	−0.52
	PBC	−1.73	−0.89	−0.71	−0.49
KS/SCAN0	2-body	−1.51	−0.34	0.09	0.20
	3-body	−0.38	−0.57	−0.64	−0.51
	4-body	0.03	0.08	0.12	0.07
	Total MBE	−1.86	−0.83	−0.43	−0.24
	PBC	−1.90	−0.76	−0.44	−0.20

Table A.5: The RPA correlation energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Input	Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
KS/PBE	2-body	−37.64	−32.22	−28.13	−22.41
	3-body	1.89	1.07	0.46	0.23
	4-body	−0.01	0.24	0.53	0.32
	Total MBE	−35.76	−30.91	−27.14	−21.86
	PBC	−35.44	−30.63	−27.14	−21.72
KS/SCAN	2-body	−36.36	−30.97	−27.27	−21.67
	3-body	0.93	0.47	0.43	0.14
	4-body	−0.05	−0.02	−0.01	0.01
	Total MBE	−35.48	−30.52	−26.85	−21.52
	PBC	−35.47	−30.48	−27.04	−21.35
KS/PBE0	2-body	−33.14	−28.95	−25.81	−20.52
	3-body	1.42	1.01	0.76	0.45
	4-body	−0.04	0.09	0.24	0.15
	Total MBE	−31.76	−27.85	−24.81	−19.92
	PBC	−31.44	−27.42	−24.51	−19.74
KS/SCAN0	2-body	−32.79	−28.49	−22.50	−20.22
	3-body	0.86	0.58	0.56	0.28
	4-body	−0.05	−0.01	−0.01	0.01
	Total MBE	−31.98	−27.92	−24.95	−19.93
	PBC	−31.74	−27.60	−24.84	−19.83

Table A.6: The MP2 correlation energies in kJ/mol obtained from the MBE and PBC approaches for all the considered systems.

Method	Ethane	Ethylene	Acetylene/I	Acetylene/II
2-body	−35.89	−35.12	−34.18	−27.46
3-body	0.68	0.83	0.98	0.63
4-body	−0.05	−0.07	−0.08	−0.04
Total MBE	−35.26	−34.36	−33.28	−26.87
PBC	−35.52	−34.18	−32.86	−26.49

A.2 Chapter 5

The effect of the examined orbitals on the RPA results for ethane, acetylene/I, and acetylene/II are summarized in Tables A.7–A.12.

Table A.7: The mean field energies of ethane in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE

Components	Orbitals	2-body	3-body	4-body	Total
DFT	KS/PBE	−5.55	6.86	−1.81	−0.50
	KS/SCAN	−8.31	−5.15	2.28	−11.18
	KS/PBE0	−3.10	3.93	−1.13	−0.30
	KS/SCAN0	−7.58	−3.44	1.41	−9.61
EXX	KS/PBE	19.45	−2.95	0.89	17.39
	KS/SCAN	18.02	−0.10	−0.23	17.69
	KS/PBE0	16.03	−1.29	0.32	15.06
	KS/SCAN0	15.76	−0.14	−0.07	15.55
RSE	KS/PBE	−4.46	1.46	−0.52	−3.52
	KS/SCAN	−3.13	−0.36	0.11	−3.38
	KS/PBE0	−1.93	0.36	−0.14	−1.71
	KS/SCAN0	−1.51	−0.38	0.03	−1.86
1RDM,quad	semi/PBE	−1.71	0.84	–	–
HF	-	12.95	−0.62	0.03	12.36

Table A.8: The mean field energies of acetylene/I in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE

Components	Orbitals	2-body	3-body	4-body	Total
DFT	KS/PBE	−15.88	6.28	−2.07	−11.67
	KS/SCAN	−21.08	−1.53	0.57	−22.04
	KS/PBE0	−14.75	3.67	−1.30	−12.38
	KS/SCAN0	−20.44	−0.92	0.29	−21.07
EXX	KS/PBE	6.53	−2.47	0.6	4.66
	KS/SCAN	3.81	−0.15	−0.26	3.40
	KS/PBE0	3.37	−1.27	0.15	2.25
	KS/SCAN0	2.32	−0.22	−0.14	1.96
RSE	KS/PBE	−3.06	0.81	−0.24	−2.49
	KS/SCAN	−0.95	−0.67	0.19	−1.43
	KS/PBE0	−0.79	0.04	0.01	−0.74
	KS/SCAN0	0.09	−0.64	0.12	−0.43
1RDM,quad	semi/PBE	−1.63	0.72	–	–
HF	-	1.64	−0.90	0.03	0.77

Table A.9: The mean field energies of acetylene/II in kJ/mol obtained from the RPA calculations using the MBE approach. The values of EXX and RSE components do not depend on the transformation from KS/PBE to semi/PBE

Components	Orbitals	2-body	3-body	4-body	Total
DFT	KS/PBE	−15.64	5.42	−1.18	−11.4
	KS/SCAN	−19.17	−0.82	0.29	−19.7
	KS/PBE0	−14.54	3.35	−0.73	−11.92
	KS/SCAN0	−18.79	−0.36	0.14	−19.01
EXX	KS/PBE	1.63	−1.36	0.41	0.68
	KS/SCAN	−0.69	0.54	−0.13	−0.28
	KS/PBE0	−0.95	−0.41	0.13	−1.23
	KS/SCAN0	−1.84	0.42	−0.05	−1.47
RSE	KS/PBE	−2.4	0.6	−0.15	−1.95
	KS/SCAN	−0.58	−0.58	0.11	−1.05
	KS/PBE0	−0.55	0.02	0.01	−0.52
	KS/SCAN0	0.2	−0.51	0.07	−0.24
1RDM,quad	semi/PBE	−1.29	0.53	–	–
HF	-	−2.24	−0.15	0.05	−2.34

Table A.10: The correlation energies of ethane in kJ/mol obtained from the RPA calculations using the MBE approach.

Components	Orbitals	2-body	3-body	4-body	Total
RPA	KS/PBE	−37.56	1.89	−0.01	−35.68
	KS/SCAN	−36.29	0.93	−0.05	−35.41
	KS/PBE0	−33.07	1.42	−0.04	−31.69
	KS/SCAN0	−32.72	0.86	−0.05	−31.91
	HF	−25.18	1.34	–	–
	semi/PBE	−28.68	0.97	–	–
SOSEX	KS/PBE	−1.51	0.45	–	–
	semi/PBE	−0.36	0.31	–	–
	HF	0.12	0.01	–	–
2g	semi/PBE	−6.37	0.51	–	–
	HF	−5.79	0.58	–	–

Table A.11: The correlation energies of acetylene/I in kJ/mol obtained from the RPA calculations using the MBE approach.

Components	Orbitals	2-body	3-body	4-body	Total
RPA	KS/PBE	−28.09	0.46	0.53	−27.10
	KS/SCAN	−27.25	0.43	−0.01	−26.83
	KS/PBE0	−25.78	0.76	0.24	−24.78
	KS/SCAN0	−25.48	0.56	−0.01	−24.93
	HF	−23.19	1.65	–	–
	semi/PBE	−23.12	0.38	–	–
SOSEX	KS/PBE	−3.31	0.68	–	–
	semi/PBE	−1.21	0.39	–	–
	HF	−0.25	−0.03	–	–
2g	semi/PBE	−3.65	0.21	–	–
	HF	−4.11	0.61	–	–

Table A.12: The correlation energies of acetylene/II in kJ/mol obtained from the RPA calculations using the MBE approach.

Components	Orbitals	2-body	3-body	4-body	Total
RPA	KS/PBE	−22.38	0.23	0.32	−21.83
	KS/SCAN	−21.64	0.14	0.01	−21.49
	KS/PBE0	−20.50	0.45	0.15	−19.90
	KS/SCAN0	−20.19	0.28	0.01	−19.90
	HF	−18.48	1.13	–	–
	semi/PBE	−18.35	0.27	–	–
SOSEX	KS/PBE	−2.87	0.59	–	–
	semi/PBE	−1.17	0.34	–	–
	HF	−0.41	0.03	–	–
2g	semi/PBE	−3.01	0.14	–	–
	HF	−3.35	0.43	–	–

A.3 Chapter 6

The results of the n-body energies and total energies of the examined methods for all the systems are listed in Tables A.13–A.16. The results of the correction scheme for ethane, acetylene/I, and acetylene/II are summarized in Tables A.17–A.19.

Table A.13: Total 2-body energies of the examined methods summed over all distances (kJ/mol) for all the considered systems.

Methods	Ethane	Ethylene	Acetylene/I	Acetylene/II
		KS/PBE		
$E_{\text{total}}^{\text{RPA}}$	−18.11	−17.94	−21.56	−20.75
$E_{\text{total}}^{\text{RPA+RSE}}$	−22.57	−21.15	−24.62	−23.16
$E_{\text{total}}^{\text{rPT2}}$	−24.08	−23.46	−27.93	−26.03
		semi/PBE		
$E_{\text{total}}^{\text{RPA+MBPT3}}$	−22.12	−21.65	−26.14	−24.60
		KS/SCAN		
$E_{\text{total}}^{\text{RPA}}$	−18.27	−18.84	−23.44	−22.33
$E_{\text{total}}^{\text{RPA+RSE}}$	−21.40	−20.31	−24.39	−22.91
		KS/PBE0		
$E_{\text{total}}^{\text{RPA}}$	−17.04	−17.65	−22.41	−21.45
$E_{\text{total}}^{\text{RPA+RSE}}$	−18.97	−18.67	−23.20	−22.00
		KS/SCAN0		
$E_{\text{total}}^{\text{RPA}}$	−16.96	−17.92	−23.16	−22.03
$E_{\text{total}}^{\text{RPA+RSE}}$	−18.47	−18.26	−23.07	−21.83
		HF		
$E_{\text{total}}^{\text{RPA}}$	−12.23	−14.74	−21.55	−20.71
$E_{\text{total}}^{\text{rPT2}}$	−12.11	−14.63	−21.80	−21.13
$E_{\text{total}}^{\text{RPA+MBPT3}}$	−17.90	−19.85	−25.91	−24.47
$E_{\text{total}}^{\text{MP2}}$	−22.80	−25.61	−32.51	−29.70
$E_{\text{total}}^{\text{MP2.5}}$	−22.40	−24.30	−29.70	−27.60
$E_{\text{total}}^{\text{MP3}}$	−21.90	−23.07	−26.93	−25.44
$E_{\text{total}}^{\text{CCSD}}$	−18.25	−18.08	−22.39	−21.59
$E_{\text{total}}^{\text{CCSD(T)}}$	−23.89	−23.49	−27.66	−25.86

Table A.14: Total 3-body energies of the examined methods summed over all distances (kJ/mol) for all the considered systems.

Methods	Ethane	Ethylene	Acetylene/I	Acetylene/II
		KS/PBE		
$E_{\text{total}}^{\text{RPA}}$	−1.06	−1.10	−2.02	−1.13
$E_{\text{total}}^{\text{RPA+RSE}}$	0.41	−0.07	−1.21	−0.53
$E_{\text{total}}^{\text{rPT2}}$	0.86	0.63	−0.53	0.06
		semi/PBE		
$E_{\text{total}}^{\text{RPA+MBPT3}}$	1.15	0.99	0.02	0.52
		KS/SCAN		
$E_{\text{total}}^{\text{RPA}}$	0.81	0.94	0.27	0.68
$E_{\text{total}}^{\text{RPA+RSE}}$	0.47	0.33	−0.39	0.11
		KS/PBE0		
$E_{\text{total}}^{\text{RPA}}$	0.12	0.23	−0.49	0.12
$E_{\text{total}}^{\text{RPA+RSE}}$	0.49	0.36	−0.45	0.07
		KS/SCAN0		
$E_{\text{total}}^{\text{RPA}}$	0.72	0.91	0.33	0.70
$E_{\text{total}}^{\text{RPA+RSE}}$	0.34	0.35	−0.30	0.19
		HF		
$E_{\text{total}}^{\text{RPA}}$	0.73	1.12	0.74	0.98
$E_{\text{total}}^{\text{rPT2}}$	0.74	1.14	0.72	1.02
$E_{\text{total}}^{\text{RPA+MBPT3}}$	1.32	1.75	1.33	1.45
$E_{\text{total}}^{\text{MP2}}$	0.06	0.45	0.08	0.48
$E_{\text{total}}^{\text{MP2.5}}$	1.06	1.59	1.31	1.41
$E_{\text{total}}^{\text{MP3}}$	2.06	2.73	2.54	2.33
$E_{\text{total}}^{\text{CCSD}}$	1.25	1.36	0.74	0.98
$E_{\text{total}}^{\text{CCSD(T)}}$	1.45	1.56	0.97	1.15

Table A.15: Total 4-body energies of the examined methods summed over all distances (kJ/mol).

Methods	Ethane	Ethylene	Acetylene/I	Acetylene/II
		KS/PBE		
$E_{\text{total}}^{\text{RPA}}$	0.88	0.86	1.13	0.73
$E_{\text{total}}^{\text{RPA+RSE}}$	0.36	0.53	0.89	0.58
		KS/SCAN		
$E_{\text{total}}^{\text{RPA}}$	−0.28	−0.33	−0.27	−0.12
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.17	−0.17	−0.08	−0.01
		KS/PBE0		
$E_{\text{total}}^{\text{RPA}}$	0.28	0.28	0.39	0.28
$E_{\text{total}}^{\text{RPA+RSE}}$	0.14	0.22	0.40	0.29
		KS/SCAN0		
$E_{\text{total}}^{\text{RPA}}$	−0.12	−0.16	−0.15	−0.04
$E_{\text{total}}^{\text{RPA+RSE}}$	−0.09	−0.08	−0.03	0.03
		HF		
$E_{\text{total}}^{\text{MP2}}$	−0.02	−0.05	−0.05	0.01
$E_{\text{total}}^{\text{CCSD}}$	−0.09	−0.10	−0.08	−0.02
$E_{\text{total}}^{\text{CCSD(T)}}$	−0.10	−0.11	−0.10	−0.03

Table A.16: Total n-body energies of the examined methods summed over all distances (kJ/mol).

Methods	Ethane	Ethylene	Acetylene/I	Acetylene/II
		KS/PBE		
$E_{\text{total}}^{\text{RPA}}$	−18.29	−18.15	−22.44	−21.16
$E_{\text{total}}^{\text{RPA+RSE}}$	−21.82	−20.67	−24.94	−23.12
		KS/SCAN		
$E_{\text{total}}^{\text{RPA}}$	−17.73	−18.23	−23.42	−21.78
$E_{\text{total}}^{\text{RPA+RSE}}$	−21.11	−20.14	−24.85	−22.83
		KS/PBE0		
$E_{\text{total}}^{\text{RPA}}$	−16.63	−17.15	−22.53	−21.13
$E_{\text{total}}^{\text{RPA+RSE}}$	−18.41	−18.13	−23.30	−21.67
		KS/SCAN0		
$E_{\text{total}}^{\text{RPA}}$	−16.36	−17.17	−22.97	−21.37
$E_{\text{total}}^{\text{RPA+RSE}}$	−18.29	−18.05	−23.42	−21.64
		HF		
$E_{\text{total}}^{\text{MP2}}$	−22.75	−25.21	−32.51	−29.22
$E_{\text{total}}^{\text{CCSD}}$	−17.17	−16.93	−21.82	−20.72
$E_{\text{total}}^{\text{CCSD(T)}}$	−22.54	−22.04	−26.77	−24.74

Table A.17: The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for ethane. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods.

Percentage	n-body	RPA+RSE				MP2
		KS/PBE	KS/SCAN	KS/PBE0	KS/SCAN0	–
100%	2-body	1.32	2.50	4.93	5.43	1.09
	3-body	−1.04	−0.97	−0.95	−1.10	−1.38
	4-body	0.45	−0.07	0.23	0.01	0.08
	Total	0.74	1.45	4.21	4.34	−0.21
90%	2-body	0.11	0.17	0.20	0.24	−0.01
	3-body	−0.37	−0.36	−0.21	−0.26	−0.06
	4-body	0.39	−0.06	0.20	0.01	0.06
	Total	0.12	−0.25	0.19	−0.02	0.00
70%	2-body	0.02	0.04	0.05	0.05	−0.01
	3-body	−0.10	−0.12	−0.07	−0.09	−0.08
	4-body	0.28	−0.04	0.14	0.01	0.04
	Total	0.20	−0.12	0.11	−0.03	−0.05
50%	2-body	0.01	0.01	0.02	0.02	0.00
	3-body	−0.04	−0.05	−0.03	−0.04	−0.04
	4-body	0.16	−0.01	0.08	0.01	0.02
	Total	0.12	−0.05	0.16	−0.01	−0.03
30%	2-body	0.00	0.00	0.01	0.01	0.00
	3-body	−0.02	−0.02	−0.01	−0.01	−0.01
	4-body	0.08	−0.01	0.04	0.00	0.01
	Total	0.07	−0.03	0.04	0.00	0.00
10%	2-body	0.00	0.00	0.00	0.00	0.00
	3-body	0.00	−0.01	0.00	0.00	−0.01
	4-body	0.02	0.00	0.01	0.00	0.00
	Total	0.01	−0.01	0.00	−0.01	−0.01

Table A.18: The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for acetylene/I. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods.

Percentage	n-body	RPA+RSE				MP2
		KS/PBE	KS/SCAN	KS/PBE0	KS/SCAN0	–
100%	2-body	3.03	3.27	4.44	4.58	−4.87
	3-body	−2.18	−1.35	−1.42	−1.27	−0.89
	4-body	0.99	0.03	0.50	0.08	0.05
	Total	1.84	1.94	3.52	3.39	−5.71
90%	2-body	0.03	0.07	0.07	0.08	−0.07
	3-body	−0.43	−0.35	−0.21	−0.23	−0.16
	4-body	0.88	0.03	0.44	0.07	0.04
	Total	0.47	−0.25	0.30	−0.08	−0.19
70%	2-body	0.00	0.02	0.02	0.02	−0.02
	3-body	−0.12	−0.11	−0.06	−0.08	−0.06
	4-body	0.63	0.05	0.31	0.06	0.03
	Total	0.51	−0.05	0.27	0.00	−0.05
50%	2-body	0.00	0.01	0.01	0.01	−0.01
	3-body	−0.05	−0.05	−0.02	−0.03	−0.01
	4-body	0.39	0.07	0.18	0.06	0.03
	Total	0.34	0.03	0.17	0.03	0.01
30%	2-body	0.00	0.00	0.01	0.00	−0.01
	3-body	−0.02	−0.02	−0.01	−0.02	−0.03
	4-body	0.18	0.05	0.08	0.03	0.00
	Total	0.16	0.03	0.07	0.02	−0.04
10%	2-body	0.00	0.00	0.00	0.00	−0.01
	3-body	−0.01	−0.01	0.00	−0.01	−0.01
	4-body	0.05	0.02	0.02	0.01	0.00
	Total	0.04	0.01	0.02	0.01	−0.01

Table A.19: The errors (in kJ/mol) of the values obtained using correction scheme with respect to the CCSD(T) benchmark for acetylene/II. Here, the percentage (%) represents the portion of the CCSD(T) results replaced by the results of approximate methods.

Percentage	n-body	RPA+RSE				MP2
		KS/PBE	KS/SCAN	KS/PBE0	KS/SCAN0	–
100%	2-body	2.71	2.94	3.86	4.02	−3.83
	3-body	−1.68	−1.05	−1.09	−0.96	−0.68
	4-body	0.60	0.01	0.30	0.04	0.04
	Total	1.62	1.90	0.08	3.10	−4.47
90%	2-body	0.05	0.06	0.07	0.07	−0.09
	3-body	−0.42	−0.36	−0.18	−0.19	−0.10
	4-body	0.54	0.01	0.27	0.04	0.04
	Total	0.17	−0.29	0.16	−0.09	−0.16
70%	2-body	0.03	0.02	0.03	0.02	−0.05
	3-body	−0.06	−0.08	−0.05	−0.06	−0.06
	4-body	0.37	0.02	0.18	0.03	0.02
	Total	0.33	−0.04	0.16	−0.01	−0.08
50%	2-body	0.03	0.01	0.02	0.01	−0.04
	3-body	−0.03	−0.04	−0.02	−0.02	0.00
	4-body	0.24	0.03	0.12	0.03	0.03
	Total	0.23	−0.01	0.11	0.01	−0.02
30%	2-body	0.01	0.00	0.01	0.00	−0.03
	3-body	0.00	−0.01	−0.01	−0.01	−0.01
	4-body	0.11	0.02	0.05	0.02	0.02
	Total	0.12	0.01	0.05	0.01	−0.03
10%	2-body	0.00	0.00	0.00	0.00	−0.01
	3-body	−0.01	−0.01	0.00	0.00	0.00
	4-body	0.03	0.01	0.01	0.00	0.00
	Total	0.02	0.00	0.01	0.00	−0.01