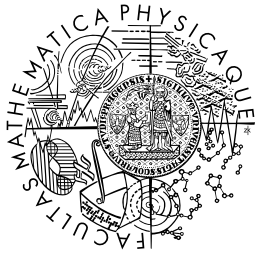

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Jaroslav Hofierka

Lattice energies of molecular solids

Department of Chemical Physics and Optics

Supervisor of the bachelor thesis: Mgr. Jiří Klimeš, Ph.D.
Study programme: Physics
Study branch: General Physics

Prague 2017

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources.

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 ..... signature of the author

Title: Lattice energies of molecular solids

Author: Jaroslav Hofierka

Department: Department of Chemical Physics and Optics

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

Abstract: Molecular solids are important materials with many applications in various fields of science and industry. They are often characterized by a rich phase diagram and the ability to adopt multiple crystal structures (polymorphism). To describe small energy differences between various phases or polymorphs, accurate quantum mechanical methods are needed. In this thesis, lattice energies of methane, methanol, ammonia, and carbon dioxide are calculated using two different approaches, namely, the fragment approach and the periodic boundary conditions (PBC) approach. These two schemes have different requirements in terms of compute cost and human time needed to obtain precise results. In the fragment scheme, the Hartree-Fock, MP2, and CCSD(T) quantum mechanical methods are employed. In the PBC scheme, the Hartree-Fock and MP2 lattice energies are calculated. For all four systems, which differ in the nature of prevalent intermolecular interactions, a very good agreement in the range of 0.1 – 0.6 kJ/mol was found between both approaches at the MP2 level.

Keywords: molecular solids, lattice energies, ab-initio methods

Title: Vazebné energie molekulárních krystalů

Autor: Jaroslav Hofierka

Katedra: Katedra chemické fyziky a optiky

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

Abstrakt: Molekulárne kryštály sú dôležité materiály s mnohými aplikáciami v rôznych vedných odboroch a priemysle. Často majú bohatý fázový diagram a môžu sa vyskytovať v rôznych kryštálových štruktúrach. K popisu malých energetických rozdielov medzi jednotlivými fázami alebo kryštálovými štruktúrami sú potrebné presné kvantovo-mechanické metódy. V tejto práci počítame väzbové energie kryštálov metánu, metanolu, amoniaku a oxidu uhličitého dvomi rôznymi prístupmi, a to fragmentovým prístupom a prístupom s periodickými okrajovými podmienkami. Tieto dve metódy majú rozličné nároky na výpočtový čas a ľudské zdroje. V rámci fragmentového prístupu sme použili kvantovo mechanické metódy Hartree-Fock (HF), MP2 a CCSD(T). S periodickými okrajovými podmienkami sme aplikovali HF a MP2 metódy. Pre všetky skúmané systémy, ktoré sa navzájom líšia dominantnými medzimolekulárnymi interakciami, sme obomi prístupmi získali zhodné výsledky s odchýlkami 0,1 – 0,6 kJ/mol na úrovni MP2 metódy.

Klíčová slova: molekulární krystaly, vazebné energie, ab-initio metody

I would like to thank my supervisor Mgr. Jiří Klimeš, Ph.D., for introducing me to the world of electronic structure modelling, for his guidance, and for many helpful comments on this thesis. Also, I am grateful to my family and my girlfriend for their support and patience.

Contents

Introduction

Molecular solids are important materials that find applications in many fields of science and technology such as solid-state chemistry, pharmaceutical engineering, Earth sciences, etc. A molecular crystal can have several different crystal-packing motifs or polymorphs. This has far-reaching consequences because polymorphs can exhibit completely different physical characteristics.[1] Furthermore, sometimes we also have to deal with multicomponent systems such as cocrystals, salts, hydrates, and solvates. For example, methane clathrate hydrates are crystalline solids with methane gas molecules trapped inside various polyhedral water cages. They have attracted attention due to their abundance in the seafloor and elsewhere and their potential to serve as an energy resource [2].

One motivation to study molecular solids theoretically is often a lack of experimental methods for studying systems under real-world conditions. For example, extremely high pressures (up to 1 TPa) can be found in the upper mantles of planets such as Neptune and Uranus, which are rich in methane, ammonia, and water. Many observable properties of these planets, such as gravitational moments and atmospheric composition, are determined by the physical and chemical properties of matter within this layer. To study these intriguing molecular solids, molecular simulations at high pressures (up to 300 GPa) and temperatures (up to 7000 K) have been performed [3].

However, even an accurate description of everyday phenomena like freezing water is still a daunting task. Water ice exhibits a rich and complex phase diagram. Besides the hexagonal Ih structure, familiar as the ice and snow found in colder regions of the Earth's surface, and the closely related cubic ice Ic, there are 15 other experimentally known structures. Using widely used theoretical methods, tiny energy differences in the order of 1 kJ/mol (or even below) between various phases are very difficult to distinguish [4, 5, 6].

In this thesis, we study molecular solids formed by molecules which are bound together by non-covalent interactions. To understand the nature of these rather weak intermolecular interactions and characterize the properties of molecular solids theoretically, methods of quantum mechanics must be employed. In fact, a range of quantum mechanical methods, which differ in compute cost and accuracy, is available. Since molecular solids are large systems, the most accurate schemes such as the coupled cluster method cannot be applied directly due to a high computational cost. Another well-known method, MP2, being less demanding and still very accurate, has been employed within periodic boundary conditions only recently [7, 8]. As a next example, the RPA scheme offers good accuracy at a modest computational cost [9]. Recent benchmark tests show that while RPA underestimates lattice energies by cca 10 – 20 %, adding singles corrections can improve the results significantly [10, 11].

The widely used DFT approximative methods are computationally cheaper than MP2 or RPA, which makes them applicable to molecular solids, but they

[1] A famous example is ritonavir: a drug for treating patients infected with HIV. Several years after it was introduced to the market, a more stable and less soluble polymorph appeared. This compromised the oral bioavailability of the drug and caused the removal of the oral capsule formulation from the market [1].

are less accurate. To obtain qualitatively good results, one needs to employ dispersion[2] corrected DFT schemes, such as "D3" of Grimme et al. [12], "XDM" of Otero-de-la Roza and Johnson [13],"TS" of Tkatchenko and Scheffler [14] and "MBD" of Ambrosetti et al. [15]. These methods have been tested on a number of molecular solids with good results, but their accuracy is not entirely satisfactory [16, 13].

In principle, empirical force fields can also be applied to lattice energy calculations. Specifically, their low computational cost makes them suitable for crystal structure prediction, where the energies of a large number of possible crystal structures need to be evaluated. For crystal structure prediction, the force fields are usually developed by calculating ab-initio interaction energies of molecular dimers and trimers and subsequent fitting by analytic functions. Another approach is to use parameters derived from experimental data [17].

The coupled cluster (CC) method, widely used for benchmark calculations because of its high accuracy, cannot be directly used for molecular solids within periodic boundary conditions due to a steep computational scaling [18]. However, coupled cluster calculations can be done using the fragment approach, decomposing the lattice energy into monomer, dimer, trimer, etc., contributions, and summing up the results to obtain the total lattice energy of the solid. The principal problem of fragment-based methods is the convergence with respect to the cluster size. This is difficult to assess because the cost of the cluster calculations increases quickly with the cluster size. Many schemes have been developed for defining how to account for the interactions between the fragments and obtaining a converged result [19]. Recently, a well-converged lattice energy of benzene was obtained by Yang et al. [20]. Apart from this direct coupled cluster approach, many hybrid models combining CC with less demanding methods have been proposed. Some researchers used the CC method to correct data obtained at a lower level of theory [21, 22]. Alternatively, a combination of CC and force field fitting for the environment can be used to yield results for molecular solids [23, 24].

In reality, it is often difficult to obtain accurate lattice energies and published results often vary by more than 20%. In order to obtain a meaningful result, one needs enough computer resources and a sufficiently accurate method. In this thesis, our goal is to calculate lattice energies of molecular solids using two different approaches, the periodic boundary conditions (PBC) and the fragment approach, and compare the results. Performing calculations using the Hartree-Fock (HF) and MP2 method, we wish to understand the convergence behavior of specific implementations of these quantum mechanical methods. We discuss how the PBC and fragment approaches differ in terms of time required to set-up and perform the calculations. Moreover, we assess their effectiveness for our selected set of four molecular solids, which differ in terms of prevalent intermolecular interactions. Besides the HF and MP2 methods, we also employ the CCSD(T) method in the fragment approach and check differences between MP2 and CCSD(T) energies in order to achieve higher accuracy. In practice, both approaches have their strengths and weaknesses and their use poses several challenges; for instance, in the PBC approach we confront an unknown basis set convergence behavior, whereas in the fragment approach we are facing the problem of estimating the

[2]Dispersion is a weak intermolecular interaction arising from quantum-induced instantaneous polarization multipoles in molecules.

proper cutoff distance to achieve the desired accuracy. At the end of this thesis, we compare our results with experimental data and some earlier published results.

To make our study as general as possible, we picked four systems with different physico-chemical characteristics: methane, methanol, ammonia, and carbon dioxide. While methane, ammonia, and carbon dioxide (familiar as dry ice) are mostly held together by dispersion forces, we consider methanol to be mostly bonded by hydrogen bonds. However, since a molecule of carbon dioxide is known to have a fairly strong quadrupole moment, small electrostatic contributions to the interaction energy such as from Coulomb quadrupole-quadrupole interaction can be observed for dry ice. On the other hand, this is not the case for methane which has a zero quadrupole moment. Furthermore, in ammonia we expect weaker hydrogen bonding between neighboring molecules than in methanol. We expect that these differences might affect the suitability of the PBC and fragment schemes for the calculations of lattice energies.

1. Ab initio methods

In this chapter we present several quantum mechanical methods used for describing many electron systems. These methods are developed from first principles, *i.e.* from the Schrödinger (or Dirac) equation, hence the term "ab initio". In modern quantum chemistry, we recognize the following three main families of ab initio methods:

1. Hartree-Fock (HF) model and post-HF methods including many-body perturbation theory, configuraton interaction (CI), coupled-cluster (CC) theory, etc. The main advantage of this family of methods is a systematically improvable accuracy limited essentially only by computational resources.

2. Density functional theory (DFT) comprising local-density approximation (LDA), generalized gradient approximation (GGA), hybrid functionals, etc. Whereas computational cost of these methods is similar to the HF method, the approximate description of electron correlation makes them often more accurate than HF, which leads to a wider applicability.

3. Quantum Monte Carlo methods including variational Monte Carlo (VMC), diffusion Monte Carlo (DMC), etc. In essence, these methods sample the electronic wave function with random positions of the electrons. This allows an accurate description of large systems but at extremely high computational cost.

In this thesis we use HF and post-HF methods which are discussed in detail in the rest of this chapter.

1.1 The Hartree-Fock model

1.1.1 Fundamental relationships

To characterize many-electron systems using the quantum theory, we are interested in finding wave functions that describe a particular state of the studied system. In most cases we seek to find the stationary states of the Hamiltonian operator $\hat{H}$ which involves information about interactions and the geometry of the system. We face the task of solving the time-independent Schrödinger equation of the form

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,. \tag{1.1}$$

This is an eigenvalue equation with E representing the energy of the system.

For many-electron systems with tens of electrons the equation 1.1 cannot be solved exactly. In fact, $|\Psi\rangle$ is a wave function for both nuclei and electrons. In many cases, the electrons move faster then nuclei by several orders. Therefore we assume that they can be treated separately. In other words, we consider that atomic nuclei are infinitely heavier than electrons and thus their kinetic energy and mutual interaction can be omitted from the Hamiltonian. Eventually we are left with the Hamiltonian describing the motion of N electrons in the field of M

point charges of the nuclei:

$$\hat{H}_{\text{el}} = -\frac{1}{2}\sum_{i=1}^{N}\boldsymbol{\nabla}_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{\hat{r}_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{\hat{r}_{ij}}. \tag{1.2}$$

Here, the first term is the operator of the kinetic energy of electrons, the second term is the Coulomb interaction between nuclei and electrons and the third term represents the repulsion of electrons. The electronic Hamiltonian appears in the equation $\hat{H}_{\text{el}}\left|\Psi_{\text{el}}\right\rangle = E_{\text{el}}\left|\Psi_{\text{el}}\right\rangle$ with the electronic wave function not depending explicitly on the coordinates of nuclei $\vec{r}_A$. The total energy of the system is then:

$$E_{\text{tot}} = E_{\text{el}} + E_{\text{core}}, \tag{1.3}$$

where E_{core} is the kinetic and potential energy of the nuclei. In the next paragraphs we only consider the electronic part of the total energy.

Writing the wave function in the form $|\Psi(\vec{r}_1, \vec{r}_2, \dots)\rangle$ is not feasible. To find approximate solutions to the Schrödinger equation, we introduce single-electron wave functions (orbitals) in order to approximate $|\Psi\rangle$. Let us define a spin orbital $\chi(\vec{x})$ as a wave function of a single electron characterizing both its spatial distribution and its spin, *i.e.*, $\vec{x} = (\vec{r}_i, \alpha)$. Having selected a set of spin orbitals and a number of electrons, $\chi_i(\vec{x}_j)$ denotes the i-th spin orbital occupied by an electron labeled as j. In the next paragraphs we will often use a shorter notation of $\chi_i(j)$.

As a central postulate for fermions we have the antisymmetry principle which states that a many-electron wave function must be antisymmetric with respect to the interchange of the coordinate $\vec{x}$ (involving space and spin) of any two electrons [25]. In the HF model, which forms the basis for many other quantum-chemical approximation methods, the antisymmetry is put into practice by approximating the ground state wave function of a N-electron system in the form of a single Slater determinant composed of N spin orbitals:

$$|\Psi_{\text{SD}}(\vec{x}_1, ..., \vec{x}_N)\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & ... & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & ... & \chi_N(\vec{x}_2) \\ ... & ... & ... & ... \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & ... & \chi_N(\vec{x}_N) \end{vmatrix}. \tag{1.4}$$

Using the permutation operator $\hat{P}$ and the antisymmetrizing operator $\hat{A}$ defined as $\hat{A} = \frac{1}{N!}\sum_p^{N-1}(-1)^p\hat{P}$, we may write:

$$|\Psi_{\text{SD}}\rangle = \frac{1}{\sqrt{N!}}\sum_{p=0}^{N-1}(-1)^p\hat{P}\,|\chi_1(\vec{x}_1)\chi_2(\vec{x}_2)...\chi_N(\vec{x}_N)\rangle \tag{1.5}$$

$$= \frac{1}{\sqrt{N!}}\left(\hat{I} - \sum_{ij}\hat{P}_{ij} + \sum_{ijk}\hat{P}_{ijk} - ...\right)|\phi\rangle = \sqrt{N!}\hat{A}\,|\phi\rangle\,, \tag{1.6}$$

where $\hat{P}_{ij}$ and $\hat{P}_{ijk}$ generate permutations of two and three electron coordinates, etc., and we introduced $|\phi\rangle$ as the product of diagonal elements of the Slater determinant.

Let us assume that the set of N spin orbitals is orthonormal. The variational principle states that the best wave function is the one that minimizes the energy

$$E_{\mathrm{SD}} = \langle \Psi_{\mathrm{SD}} | \hat{H}_{\mathrm{el}} | \Psi_{\mathrm{SD}} \rangle \,. \tag{1.7}$$

Since $\hat{A} = \hat{A}^{\dagger}$ and $\left[\hat{A}, \hat{H}\right] = 0$ and $\hat{A}\hat{A} = \hat{A}$ [1] we may substitute from Eq. 1.6 and simplify this as follows:

$$E_{\mathrm{SD}} = N! \, \langle \hat{A}\phi | \hat{H}_{el} | \hat{A}\phi \rangle = \sum_{p} (-1)^{p} \, \langle \phi | \hat{H}_{el} | \hat{P}\phi \rangle \,. \tag{1.8}$$

Let us now derive an expression for the energy of a single Slater determinant. First, recalling the equation 1.2, we introduce the following notation:

$$\hat{H}_{\mathrm{el}} = \sum_{i=1}^{N} \hat{h}_i + \sum_{i=1}^{N} \sum_{j>i}^{N} \hat{g}_{ij}, \tag{1.9}$$

where $\hat{h}_i$ is the operator of both the kinetic energy of electrons and the Coulomb interaction between nuclei and electrons and $\hat{g}_{ij}$ represents the repulsion of electrons. Then for the one-particle operator $\hat{h}_i$ we can write:

$$\begin{aligned} \langle \phi | \hat{h}_i | \phi \rangle &= \langle \chi_1(1)\chi_2(2)...\chi_N(N) | \hat{h}_i | \chi_1(1)\chi_2(2)...\chi_N(N) \rangle & (1.10) \\ &= \langle \chi_1(1)|\chi_1(1)\rangle \, \langle \chi_2(2)|\chi_2(2)\rangle \, ... \, \langle \chi_i(i)|\chi_i(i)\rangle \, ... \, \langle \chi_N(N)|\chi_N(N)\rangle & (1.11) \\ &= \langle \chi_i(i) | \hat{h}_i | \chi_i(i) \rangle = h_i, \;\; i = 1, 2, 3 \ldots N. & (1.12) \end{aligned}$$

It follows from the above equation and the orthonormality of the spin orbitals that all matrix elements of $\hat{h}_i$ involving a permutation operator are zero. For the two-electron operator $\hat{g}_{ij}$, only the identity and $\hat{P}_{ij}$ operators can give non-zero contributions. A three-electron (and four-, five-, etc.) permutation will always be zero. The term arising from the identity operator can be simplified:

$$\begin{aligned} \langle \phi | \hat{g}_{ij} | \phi \rangle &= \langle \chi_1(1)|\chi_1(1)\rangle \, ... \, \langle \chi_i(i)\chi_j(j) | \hat{g}_{ij} | \chi_i(i)\chi_j(j)\rangle \, ... \, \langle \chi_N(N)|\chi_N(N)\rangle & (1.13) \\ &= \langle \chi_i(i)\chi_j(j) | \hat{g}_{ij} | \chi_i(i)\chi_j(j) \rangle = J_{ij}. & (1.14) \end{aligned}$$

Physically, J_{ij} is a matrix element representing the coulomb interaction between two electrons. Finally we have:

$$\begin{aligned} \left\langle \phi \middle| \hat{g}_{ij} \middle| \hat{P}_{ij}\phi \right\rangle &= \langle \chi_i(i)\chi_j(j)...\chi_N | \hat{g}_{ij} | \chi_j(i)\chi_i(j)...\chi_N(N) \rangle & (1.15) \\ &= \langle \chi_i(i)\chi_j(j) | \hat{g}_{ij} | \chi_j(i)\chi_i(j) \rangle = K_{ij}. & (1.16) \end{aligned}$$

The K_{ij} matrix element is called the exchange integral, and has no classical analogy. Using these results we can express the energy of a Slater determinant as follows:

$$E_{\mathrm{SD}} = \sum_{i=1}^{N} h_i + \sum_{i=1}^{N} \sum_{j>i}^{N} (J_{ij} - K_{ij}). \tag{1.17}$$

[1] The idempotency of $\hat{A}$ is clear since a repeated action of $\hat{A}$ creates $N!$ terms out of a single term and there are only $N!$ unique permutations in a group of N elements.

1.1.2 The HF and Roothaan equations

By minimizing the functional $E_{\mathrm{SD}}[\chi_i]$ with respect to the spin orbitals we can derive the Hartree-Fock equations which help us determine the optimal choice of the spin orbitals (and, consequently, optimal energy value).

Let us define the Coulomb operator $\hat{J}$ and the exchange operator $\hat{K}$ in this way:

$$\hat{J}_i \left|\chi_j(2)\right\rangle = \left\langle\chi_i(1)\right|\hat{g}_{12}\left|\chi_i(1)\right\rangle \left|\chi_j(2)\right\rangle \tag{1.18}$$

$$\hat{K}_i \left|\chi_j(2)\right\rangle = \left\langle\chi_i(1)\right|\hat{g}_{12}\left|\chi_j(1)\right\rangle \left|\chi_i(2)\right\rangle . \tag{1.19}$$

Then, the Fock operator has the form

$$\hat{F}_i = \hat{h}_i + \sum_j^N (\hat{J}_j - \hat{K}_j). \tag{1.20}$$

Applying the method of Lagrange multipliers with the constraint $\langle\chi_i|\chi_j\rangle = \delta_{ij}$ we can write the variational problem as follows:

$$\delta E_{\mathrm{SD}} - \delta \sum_{i=1}^N \sum_{j=1}^N \left[\epsilon_{ij} \left(\langle\chi_i|\chi_j\rangle - \delta_{ij}\right)\right] = 0, \tag{1.21}$$

where ϵ_{ij} is a set of Lagrange multipliers.

Making use of the Fock operator we may rewrite the equation 1.21 as follows:

$$\sum_{i=1}^N \Big(\langle\delta\chi_i|\hat{F}_i|\chi_i\rangle + \langle\chi_i|\hat{F}_i|\delta\chi_i\rangle\Big) - \sum_{ij}^N \epsilon_{ij} \left(\langle\delta\chi_i|\chi_j\rangle + \langle\chi_i|\delta\chi_j\rangle\right) = 0 \tag{1.22}$$

$$\sum_{i=1}^N \langle\delta\chi_i|\hat{F}_i|\chi_i\rangle - \sum_{ij}^N \epsilon_{ij} \langle\delta\chi_i|\chi_j\rangle + \sum_{i=1}^N \langle\delta\chi_i|\hat{F}_i|\chi_i\rangle^* - \sum_{ij}^N \epsilon_{ij} \langle\delta\chi_j|\chi_i\rangle^* = 0, \tag{1.23}$$

where we used the fact that $\langle\chi|\hat{F}|\delta\chi\rangle^* = \langle\delta\chi|\hat{F}|\chi\rangle$ and $\langle\chi|\delta\chi\rangle^* = \langle\delta\chi|\chi\rangle$. The first two terms in eq. 1.23 must cancel, and the last two terms must cancel. Taking the complex conjugate of the last two terms and subtracting them from the first two gives $\sum_{ij}^N (\epsilon_{ij} - \epsilon_{ij}^*) \langle\delta\chi_i|\chi_j\rangle = 0$. This means that the Lagrange multipliers are elements of a Hermitian matrix. The final set of Hartree–Fock equations may be written as follows:

$$\hat{F}_i\chi_i = \sum_j^N \epsilon_{ij}\chi_j. \tag{1.24}$$

The equations may be simplified by choosing a unitary transformation that makes the matrix of Lagrange multipliers diagonal. Then we simply have:

$$\hat{F}_i\chi_i = \epsilon_i\chi_i, \quad i = 1, 2, ..., N. \tag{1.25}$$

The unique set of spin orbitals that constitute a solution to this eigenvalue equation is called the canonical set [25].

For practical use, the spin orbitals are transformed into spatial orbitals by integrating out spin. Here we only consider closed-shell calculations, in which we assume that N electrons are paired in such a way that $N/2$ spatial orbitals are doubly occupied[2]: $|\Psi_{SD}\rangle = \left|\psi_1\bar{\psi}_1 ... \psi_{N/2}\bar{\psi}_{N/2}\right\rangle$. The bar indicates that the orbital is occupied with an electron with a spin down. Moreover, in the HF method the exchange interaction K_{ij} takes place only between electrons of parallel spin. This leads to the closed-shell HF energy:

$$E_{\mathrm{SD}}^{\mathrm{CS}} = 2\sum_{i=1}^{N/2} h_i + \sum_{i=1}^{N/2} J_{ii} + \sum_{i=1}^{N/2}\sum_{j>i}^{N/2}(2J_{ij} - K_{ij}). \tag{1.26}$$

Now, the summation runs over orbitals instead of electrons.

The HF equations are integro-differential equations which are very hard to solve. In order to get results we use a basis set expansion to express the unknown spatial orbitals in terms of some pre-defined functions. Let us introduce a set of K known functions $\varphi_\mu(\vec{r})$. Then we write

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i}\varphi_\mu, \;\; i = 1, 2, ..., K. \tag{1.27}$$

In the limit of a complete basis set (infinite number of basis functions), the results reach the exact Hartree–Fock energy, which is denoted as the Hartree–Fock limit. The issue of convergence will be discussed in a succeeding chapter. Substituting Eq. 1.27 into 1.25 with spatial orbitals we get:

$$\hat{F}(\vec{r})\sum_{\nu=1}^{K} c_{\nu i}\varphi_\nu(\vec{r}) = \epsilon_i \sum_{\nu=1}^{K} c_{\nu i}\varphi_\nu(\vec{r}), \;\; i = 1, 2, ..., K \tag{1.28}$$

Multiplying by φ_μ^* from the left leads to:

$$\sum_{\nu=1}^{K} c_{\nu i}\varphi_\mu^*(\vec{r})\hat{F}(\vec{r})\varphi_\nu(\vec{r}) = \epsilon_i \sum_{\nu=1}^{K} c_{\nu i}\varphi_\mu^*(\vec{r})\varphi_\nu(\vec{r}), \;\; i = 1, 2, ..., K. \tag{1.29}$$

Integrating both sides gives the Fock matrix $\boldsymbol{F}$ and the overlap matrix $\boldsymbol{S}$. We thus obtain the Roothaan equations:

$$\sum_{\nu=1}^{K} F_{\mu\nu}c_{\nu i} = \epsilon_i \sum_{\nu=1}^{K} S_{\mu\nu}c_{\nu i}, \;\; i = 1, 2, ..., K. \tag{1.30}$$

Moreover, if we introduce the diagonal matrix $\boldsymbol{\epsilon}$ with ϵ_i on the diagonal and the $K \times K$ matrix $\boldsymbol{C}$ with the i-th column of $\boldsymbol{C}$ describing the spatial orbital ψ_i, we can, finally, write the closed-shell HF equations in compact matrix form, where all matrices are square and have as many rows and columns as there are basis functions:

$$\boldsymbol{FC} = \boldsymbol{SC\epsilon}. \tag{1.31}$$

If $\boldsymbol{S} \neq \boldsymbol{I}$, which is equivalent to the non-orthogonality of our basis functions, then this is a generalized (or pseudo) eigenvalue problem and the usual approach is to orthogonalize the basis functions, thereby reducing it to a regular eigenvalue problem.

[2]This is commonly referred to as the restricted HF (RHF) method.

1.1.3 The Self-Consistent Field (SCF) Procedure

The Roothaan equations must be solved iteratively since the Fock matrix depends on its own solutions. A very simplified HF-SCF procedure contains the following steps:

- Specify the positions and charges of nuclei.
- Choose basis functions and calculate all required integrals.
- Diagonalize the overlap matrix.
- Obtain a guess for the orbital expansion coefficients forming the density matrix.
- Calculate two-electron integrals and diagonalize the Fock matrix.
- Create a new density matrix and compare with the old density matrix. If it does not agree within a specified convergence criterion, continue the SCF cycle by returning to the previous step.

In practice, the SCF procedure with the RHF type wave functions as described above does not always provide meaningful results. For example, the constraint of doubly occupied spinorbitals in the RHF theory is inconsistent with breaking bonds to produce radicals, therefore dissociation problems cannot be handled within the RHF (closed-shell) theory.

One of the bottlenecks of a HF calculation is evaluating the two-electron integrals. With K being the number of basis functions, there are, in fact, about K^4 two-electron integrals to calculate. However, only relatively few two-electron integrals differ from zero by a non-negligible amount. The trick is, thus, to compute only those that are estimated to be larger than a given cutoff value (which needs to be chosen according to desired accuracy).

Another bottleneck is the repeated diagonalization of the Fock matrix. However, since we are usually interested mainly in the occupied orbitals, there is no need to calculate hundreds or thousands of virtual orbitals. Therefore, many specific iterative matrix diagonalization procedures, which work only with the subspace of interest and produce the lowest eigenvalues and vectors in the end, have been developed for this purpose [26].

Unfortunately, the SCF procedure described above is not guaranteed to converge. Whether it does or not might depend on the initial guess and if it converges, the quality of the initial guess obviously determines how many iterations are needed. Several iteration techniques designed for accelerating the convergence of SCF methods have been developed; let us mention only the most widely used DIIS (direct inversion of the iterative subspace) method developed by Pulay in 1980.

1.1.4 Periodic systems

For periodic systems, a specific theoretical description is needed due to their unique properties. Molecular solids contain an infinite number of electrons and nuclei and ultimately we must find an elegant way to characterize such systems

and find their properties after a finite amount of computer time. A crystal can be described as a specific atomic (or molecular) arrangement repeated all over space. This repetition is characterized by lattice vectors, which can be formed as a linear combination of the so-called unit-cell vectors:

$$\vec{R} = \sum_{i=1}^{3} \xi_i \vec{a}_i, \text{with} \, \xi_i \in \mathbb{Z}. \tag{1.32}$$

Similarly we define three vectors called reciprocal lattice vectors $\vec{b}_i$ that satisfy:

$$\vec{a}_i \vec{b}_j = 2\pi\delta_{ij}. \tag{1.33}$$

Then we can construct the reciprocal lattice analogously to the direct lattice in the Eq. 1.32 as follows:

$$\vec{G} = \sum_{i=1}^{3} \zeta_i \vec{b}_i, \;\; \zeta_i \in \mathbb{Z}. \tag{1.34}$$

The $\vec{a}_i$ vectors are not unique and we usually use a specific primitive cell called the Wigner-Seitz cell with its equivalent in reciprocal space known as the first Brillouin zone. If we denote the volume of the unit cell as $V_0 = \vec{a}_1(\vec{a}_2 \times \vec{a}_3)$ then the volume of the first Brillouin zone is:

$$V_{\text{BZ}} = \frac{(2\pi)^3}{V_0}. \tag{1.35}$$

To describe quantum-mechanically infinite periodic lattices, we are faced with the problem of normalization of a wave function, which could possibly extend over the whole space. Therefore we introduce a normalization volume Ω, which is finite for all mathematical purposes, but large enough to appear infinite for all physical considerations, i.e., a translation of the entire system by a lattice vector does not change the expectation values of quantum-mechanical observables.

Let us choose $\Omega = l^3 V_0$ with $l \gg 1$. All points that are outside Ω are mapped onto a point inside Ω through periodic boundary conditions. In particular, for any single-particle spatial orbital, we require:

$$\psi(\vec{r} + l\vec{a}_i) = \psi(\vec{r}). \tag{1.36}$$

Next, consider a single plane wave with a wave-vector $\vec{k}$ as a general case for a wave function that is not lattice periodic. We have:

$$\psi(\vec{r}) = e^{i\vec{k}\vec{r}} = e^{i\vec{k}\vec{r}} e^{il\vec{k}\vec{a}_i} = \psi(\vec{r} + l\vec{a}_i), \tag{1.37}$$

which means that $l\vec{k}\vec{a}_i = 2\pi\gamma_i$, where γ_i is an integer. The vector $\vec{k}$ is a reciprocal vector and can be expressed in terms of reciprocal basis vectors as: $\vec{k} = \sum_{i=1}^{3} \kappa_i \vec{b}_i$, where $\kappa_i \in \mathbb{R}_0$. Using these results together with the equation 1.33 leads to:

$$2\pi\gamma_j = l \sum_{i=1}^{3} \kappa_i \vec{b}_i \cdot \vec{a}_j = l \sum_{i=1}^{3} \kappa_i 2\pi\delta_{ij} = 2\pi l\kappa_j \Rightarrow \kappa_j = \frac{\gamma_j}{l}, \tag{1.38}$$

which means that imposing periodic boundary conditions places restrictions on the discrete values that the components of $\vec{k}$ can assume. As a consequence, the reciprocal lattice spanned by the vectors $\{\vec{G}\}$ is further subdivided into a k-mesh with l^3 points in the first Brillouin zone (See Fig. 1.1).

The periodicity of a crystal can be formalized in terms of translational symmetry. Quantum-mechanically, we require that all expectation values are invariant with respect to a translation by any $\vec{R}$. Let us now introduce a translation operator by defining its action on a single-particle spatial orbital as:

$$\hat{T}_R\psi(\vec{r}) = \psi(\vec{r}+\vec{R}). \tag{1.39}$$

From the translational invariance it can be shown [3] that the eigenvalues of $\hat{T}_R$ are:

$$\hat{T}_R\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\vec{R}}\psi_{\vec{k}}(\vec{r}), \tag{1.41}$$

where we introduced the quantum number $\vec{k}$. Importantly, we have all eigenvalues and eigenfunctions by considering only $\vec{k}$ from the first BZ. [4] The eigenfunctions of $\hat{T}_R$ are given by the Bloch's theorem:

Any single-particle wave function in a periodic system can be written as the product of a plane wave and a lattice-periodic part $\phi_{\vec{k}}(\vec{r}) = \phi_{\vec{k}}(\vec{r}+\vec{R})$.

Mathematically:

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{\Omega}}e^{i\vec{k}\vec{r}}\phi_{\vec{k}}(\vec{r}). \tag{1.42}$$

It is straightforward to show that for the single particle Hamiltonian in a periodic system we have $[\hat{T}_{\vec{R}}, \hat{H}] = 0$. Consequently, we can find the simultaneous eigenfunctions characterized by periodicity and energy. Besides the quantum number $\vec{k}$, we add the energy band index j and we have:

$$\hat{H}\left|\psi_{\vec{k},j}\right\rangle = \epsilon_{\vec{k},j}\left|\psi_{\vec{k},j}\right\rangle. \tag{1.43}$$

If the lattice exhibits symmetries in addition to translation then the band structure $\epsilon_j(\vec{k})$ will reflect these. In other words, the more symmetric lattice we have, the smaller portion of the Brillouin zone contains all information about the properties of the system. [5]

For practical calculations we introduce plane waves as basis functions for the expansion of our unknown Bloch waves. The expansion is of the form:

$$\psi_{\vec{k},j}(\vec{r}) = \sum_{\vec{G}} C_{\vec{G}j}(\vec{k})\chi_{\vec{k},\vec{G}}(\vec{r}) = \sum_{\vec{G}} C_{\vec{G}j}\frac{1}{\sqrt{\Omega}}e^{i(\vec{k}+\vec{G})\vec{r}}, \tag{1.44}$$

where the sums involve all $\vec{G}$ such that $\frac{\hbar^2}{2m}(\vec{k}+\vec{G})^2 \leq E_{\text{cutoff}}$, *i.e.* only plane waves up to some defined energy are included in the basis. Substituting this expression

[3] To determine the eigenvalues of this operator, we consider the probability of finding a single particle at point $\vec{r}$ which must satisfy:

$$|\psi(\vec{r})|^2 = |\psi(\vec{r}+\vec{R})|^2. \tag{1.40}$$

Thus we can write $\psi(\vec{r}+\vec{R}) = \xi(\vec{R})\psi(\vec{r})$ with $|\xi(\vec{R})|^2 = 1$ and also $\xi(\vec{R}+\vec{R}') = \xi(\vec{R})\xi(\vec{R}')$.

[4] Observe that $e^{i(\vec{k}+\vec{G})\vec{R}} = e^{i\vec{k}\vec{R}}e^{i\vec{G}\vec{R}} = e^{i\vec{k}\vec{R}}$ since $e^{i\vec{G}\vec{R}} = 1$.

[5] This part of the Brillouin zone is called the irreducible wedge.

into the HF equations (Eq. 1.25) and multiplying from the left by $\chi^*_{\vec{k},\vec{G}'}(\vec{r})$ leads to an equation:

$$\sum_{\vec{G}} C_{\vec{G}j}(\vec{k}) \int_\Omega \mathrm{d}\vec{r} e^{-i(\vec{k}+\vec{G}')\vec{r}} \hat{F}(\vec{r}) e^{i(\vec{k}+\vec{G})\vec{r}} = \epsilon_{\vec{k},j} \sum_{\vec{G}} C_{\vec{G}j}(\vec{k}) \int_\Omega \mathrm{d}\vec{r} e^{-i(\vec{k}+\vec{G}')\vec{r}} e^{i(\vec{k}+\vec{G})\vec{r}} \tag{1.45}$$

Integrating over Ω, we can define the Fock matrix $\boldsymbol{F}$ on the left hand side whereas on the right hand side we can compute the integral for every given $\vec{G}'$ and $\vec{G}$ and eventually we can write:

$$\sum_{\vec{G}} C_{\vec{G}j}(\vec{k}) F_{\vec{G},\vec{G}'}(\vec{k}) = \epsilon_{\vec{k},j} \sum_{\vec{G}} C_{\vec{G}j}(\vec{k}) \delta(\vec{G} - \vec{G}'). \tag{1.46}$$

In matrix notation we have the Roothaan equations for a periodic system with orthogonal basis functions quite analogously to the Eq. 1.31:

$$\boldsymbol{F}(\vec{k})\boldsymbol{C}(\vec{k}) = \boldsymbol{C}(\vec{k})\boldsymbol{\epsilon}(\vec{k}). \tag{1.47}$$

These matrices have a dimension equal to the number of basis functions and depend on the vector $\vec{k}$.

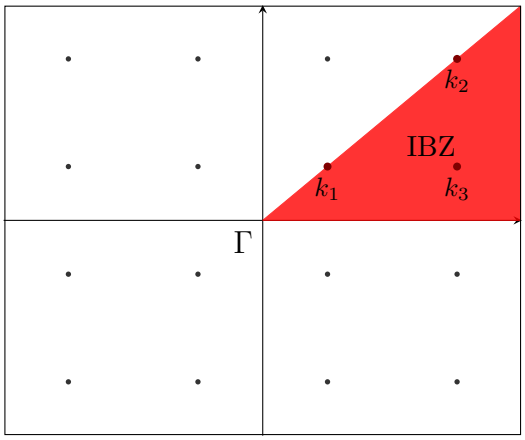

Figure 1.1: An example of the irreducible Brillouin zone (IBZ).

The evaluation of many key quantities, e.g. charge density, density-of-states, and total energy requires integration over all vectors $\vec{k}$ in the first BZ. To obtain an expression for the HF energy, it is convenient to define the density matrix as follows:

$$D_{\nu\mu}(\vec{k}) = 2 \sum_{j=1}^{N/2} C_{\nu j}(\vec{k}) C^*_{\mu j}(\vec{k}), \tag{1.48}$$

where N is the number of electrons per unit cell. Then it can be shown that the Hartree-Fock energy (without the core contribution) is given by:

$$E_{HF} = \frac{1}{2} \sum_{\mu\nu} \frac{1}{V_{\mathrm{BZ}}} \int_{\mathrm{BZ}} \mathrm{d}\vec{k} D_{\nu\mu}(\vec{k}) F_{\mu\nu}(\vec{k}) = \frac{V_0}{(2\pi)^3} \int_{\mathrm{BZ}} \mathrm{d}\vec{k} \frac{1}{2} \mathrm{Tr}[\boldsymbol{D}(\vec{k})\boldsymbol{F}(\vec{k})]. \tag{1.49}$$

Exploiting the fact that the wave functions at k-points that are close together will be almost identical, one may approximate the integration over $\vec{k}$ by a weighted

sum over a discrete set of points. Several schemes have been proposed for this k-point sampling, i.e., for selecting special points in the first BZ and for determining their relative weights in a subsequent summation.

Let us consider the Monkhorst-Pack method in which k-points in the Brillouin zone are equally spaced. Consider, for example, a two-dimensional lattice with 16 k-points as in Fig. 1.1. There are only 3 inequivalent k-points which form the irreducible Brillouin zone. Weighing each point according to the symmetry, we have:

$$\frac{1}{V_{BZ}} \int_{BZ} F(\vec{k}) \mathrm{d}\vec{k} \approx \frac{1}{4} F(\vec{k}_1) + \frac{1}{4} F(\vec{k}_2) + \frac{1}{2} F(\vec{k}_3) \tag{1.50}$$

1.2 Post-Hartree-Fock methods

The HF method is a simple mean field approach, where electron correlations are included only approximately for electrons with the same spin, whereas for electrons with opposite spin they are completely neglected. This also means that dispersion interactions are not described at all within the HF approach. To describe the important dispersion interactions, so-called post-Hartree-Fock methods based on perturbation theory are usually used.

By replacing spin orbitals that are occupied in the HF determinant by spin orbitals that are unoccupied, a whole series of determinants may be generated. These can be denoted according to how many occupied HF spin orbitals have been replaced by unoccupied spin orbitals, *i.e.*, Slater determinants that are singly, doubly, triply, quadruply, etc., excited relative to the HF determinant, up to a maximum of N excited electrons. These determinants are often referred to as Singles (S), Doubles (D), Triples (T), Quadruples (Q), etc. The total number of determinants that can be generated depends on the size of the basis set: the larger the basis, the more virtual spin orbitals, and the more determinants with a given number of excited electrons can be constructed. As shown in the Fig. 1.2, the electron correlation methods are thus two-dimensional: the larger the one-electron expansion (basis set size) and the larger the many-electron expansion (number of excitations), the better the results. In the limit of exciting all possible electrons to all possible virtual orbitals we speak of a full configuration interaction (FCI) method. The underlying concepts of the post-HF methods are briefly discussed at the end of the next subchapter.

1.2.1 Many-body perturbation theory

Rayleigh–Schrödinger perturbation theory

In this section we provide the standard derivation of the Rayleigh–Schrödinger perturbation theory. Suppose we have a Hamiltonian operator that consists of two parts, a reference $\hat{H}_0$ and a perturbation ($\hat{V}$), where the $\hat{V}$ operator is "small" in some sense compared with $\hat{H}_0$.

$$\hat{H} \left|\Phi_i\right\rangle = (\hat{H}_0 + \lambda \hat{V}) \left|\Phi_i\right\rangle = E_i \left|\Phi_i\right\rangle, \tag{1.51}$$

with λ being a parameter which will later be set to one. We will only consider cases where the perturbation is time-independent, and the reference wave function

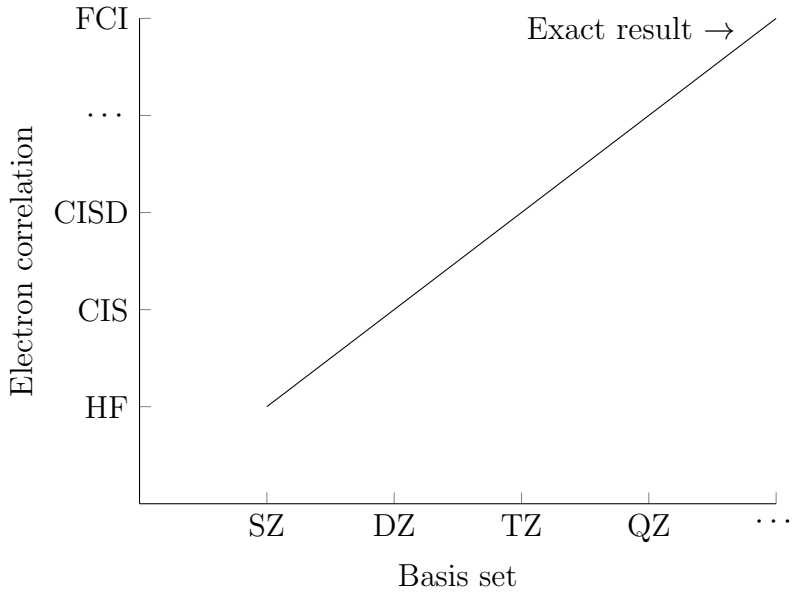

Figure 1.2: Convergence of the electron correlation methods

is non-degenerate. To keep the notation simple, we will furthermore only consider the lowest energy state. Let us assume that the Schrödinger equation for the reference Hamiltonian operator is solved:

$$\hat{H}_0 \left|\Psi_i^0\right\rangle = E_i^0 \left|\Psi_i^0\right\rangle . \tag{1.52}$$

We expand the exact eigenvalues and eigenfunctions in a Taylor series in λ:

$$E_i = E_i^0 + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + ... \tag{1.53}$$

$$|\Phi_i\rangle = \left|\Psi_i^0\right\rangle + \lambda \left|\Psi_i^{(1)}\right\rangle + \lambda^2 \left|\Psi_i^{(2)}\right\rangle + ... \tag{1.54}$$

Let us suppose that $\langle\Psi_i^0|\Phi_i\rangle = 1$ and multiply the equation 1.54 by $\langle\Psi_i^0|$. This has the consequence that all correction terms are orthogonal to the reference wave function:

$$\langle\Psi_i^0|\Psi_i^{(n)}\rangle \text{ for } n = 1, 2, 3, ... \tag{1.55}$$

Substituting Eqs. 1.53 and 1.54 into Eq. 1.51 and equating coefficients of λ^n for both sides we obtain a set of equations:

$$\hat{H}_0 \left|\Psi_i^0\right\rangle = E_i^0 \left|\Psi_i^0\right\rangle \tag{1.56}$$

$$\hat{H}_0 \left|\Psi_i^{(1)}\right\rangle + \hat{V} \left|\Psi_i^0\right\rangle = E_i^0 \left|\Psi_i^{(1)}\right\rangle + E_i^{(1)} \left|\Psi_i^0\right\rangle \tag{1.57}$$

$$\hat{H}_0 \left|\Psi_i^{(2)}\right\rangle + \hat{V} \left|\Psi_i^{(1)}\right\rangle = E_i^0 \left|\Psi_i^{(2)}\right\rangle + E_i^{(1)} \left|\Psi_i^{(1)}\right\rangle + E_i^{(2)} \left|\Psi_i^0\right\rangle \tag{1.58}$$

$$\vdots$$

Multiplying each of these equations by $\langle\Psi_i^0|$ and using the orthogonality relation

established earlier, we obtain the following expressions for the n-th order energies:

$$E_i^0 = \left\langle \Psi_i^0 \middle| \hat{H}_0 \middle| \Psi_i^0 \right\rangle \tag{1.59}$$

$$E_i^{(1)} = \langle \Psi_i^0 | \hat{V} | \Psi_i^0 \rangle \tag{1.60}$$

$$E_i^{(2)} = \langle \Psi_i^0 | \hat{V} | \Psi_i^{(1)} \rangle \tag{1.61}$$

$$\vdots$$

Let us now solve the set of equations for $|\Psi_i^{(n)}\rangle$ and consequently determine the $(n+1)$-th order correction of energy. Rearranging Eq. 1.57 we have:

$$(E_i^0 - \hat{H}_0) \left|\Psi_i^{(1)}\right\rangle = (\hat{V} - E_i^{(1)}) \left|\Psi_i^0\right\rangle . \tag{1.62}$$

Using the orthonormality relation $\langle \Psi_n^0 | \Psi_i^0 \rangle = \delta_{ni}$ and multiplying this equation by $\langle \Psi_n^0 |$ from the left, we can plug this into the expression for the second order correction of energy (Eq. 1.61) and obtain:

$$E_i^{(2)} = \langle \Psi_i^0 | \hat{V} | \Psi_i^{(1)} \rangle = \sum_n \langle \Psi_i^0 | \hat{V} | \Psi_n^0 \rangle \langle \Psi_n^0 | \Psi_i^{(1)} \rangle = \sum_{n \neq i} \frac{| \langle \Psi_i^0 | \hat{V} | \Psi_n^0 \rangle |^2}{E_i^0 - E_n^0}. \tag{1.63}$$

The higher order terms can be obtained analogously.

Møller-Plesset perturbation theory

If we take the perturbed Hamiltonian as a sum over Fock operators, we enter the Møller–Plesset (MP) perturbation theory. The sum of Fock operators counts the average electron–electron repulsion twice, and the perturbation becomes:

$$\hat{V} = \sum_i^N \sum_{j>i}^N \hat{g}_{ij} - \sum_i^N \sum_j^N \langle \hat{g}_{ij} \rangle . \tag{1.64}$$

The zeroth-order wave function is the HF determinant, and the zeroth-order energy is just a sum of orbital energies:

$$\langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle = E_0^0 = \sum_i^N \epsilon_i. \tag{1.65}$$

The Hartree-Fock energy is the sum of zeroth and first-order energies and the first correction occurs in the second order of the perturbation theory. Inspecting the general result for the second-order energy obtained in the last section (Eq. 1.63), we see that we first need to determine the states $|\Psi_n^0\rangle$. Singly excited Slater determinants do not contribute because of the Brillouin theorem [6] and moreover, triply excited states do not mix with $|\Psi_0\rangle$ because of the two-particle nature of the perturbation. Thus we only consider doubly excited Slater determinants of the form $|\Psi_{ij}^{ab}\rangle$, with two electrons excited from occupied orbitals i and j to virtual

[6] The theorem states that $\langle \psi_0 | \hat{H} | \psi_a^r \rangle = 0$ which is simply an off-diagonal element of the Fock matrix $\langle \chi_a | \hat{F} | \chi_r \rangle$. But the SCF procedure's goal was to diagonalize the Fock matrix and hence for an optimized wavefunction this quantity must be zero.

orbitals a and b. For this Slater determinant we have the eigenvalue equation of the form:

$$\hat{H}_0 \left|\Psi_{ij}^{ab}\right\rangle = \left(E_0^0 - (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)\right)\left|\Psi_{ij}^{ab}\right\rangle . \tag{1.66}$$

Therefore we have for the second-order correction:

$$E_0^{(2)} = \sum_{j>i,\, b>a} \frac{\left|\left\langle\Psi_0\right|\sum_{s>r}\hat{g}_{sr}\left|\Psi_{ij}^{ab}\right\rangle\right|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} = \sum_{j>i,\, b>a} \frac{\left|\left\langle ij|ab\right\rangle - \left\langle ij|ba\right\rangle\right|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \tag{1.67}$$

with the sums running over all occupied and virtual orbitals without double-counting and we make use of this notation for two-electron integrals:

$$\langle ij|ab\rangle = \int \chi_i^*(1)\chi_j^*(2)\frac{1}{|\vec{r}_1 - \vec{r}_2|}\chi_a(1)\chi_b(2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2. \tag{1.68}$$

The third-order correction is also calculated using the doubly excited determinants. Triply excited Slater determinants only arise in the fourth-order correction (MP4 model). In general, the more poorly the HF wave function describes the system, the larger are the correction terms, and the more terms must be included to achieve a given level of accuracy. If the reference state is a poor description of the system, the convergence may be so slow that perturbation methods cannot be used. Moreover, in the MP theory there is no guarantee that the energy will be lower than the exact energy. This is in contrast to some other methods, such as the configuration interaction (CI) method[7], where the the energy is determined variationally, and is thus an upper bound to the exact energy. However, the absence of a variational bound can allow for error cancellations. Furthermore, the MP theory, unlike CI, is a size-extensive method, which is a desirable feature especially when studying large systems.

1.2.2 Coupled clusters

Whereas perturbation methods add all types of corrections (S, D, T, etc.) to the reference wave function to a given order (2, 3, 4, etc.), the idea in Coupled Cluster (CC) methods is to include all corrections of a given type to infinite order. This method also has the features of size-extensivity and size-consistency.[8] Let us start by defining an excitation operator $\hat{T}$ as follows:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ... + \hat{T}_N, \tag{1.69}$$

where the $\hat{T}_i$ operator acting on the reference (HF Slater) determinant $|\Psi_0\rangle$ generates all i-th excited Slater determinants, for example:

$$\hat{T}_1\left|\Psi_0\right\rangle = \sum_i^{N_{occ}}\sum_r^{N_{virt}} t_i^r \left|\Psi_i^r\right\rangle , \tag{1.70}$$

[7] In CI, we approximate the exact wave function as a linear combination of Slater determinants $\Psi_{\mathrm{CI}} = c_0\Psi_{SD} + \sum_S c_S\Psi_S + \sum_D c_D\Psi_D + ..$ where subscripts S, D, etc., indicate determinants that are singly, doubly, etc., excited relative to the HF configuration. The expansion coefficients are determined by requiring that the energy is a minimum.

[8] The size-extensivity refers to the linear scaling of a method with the number of electrons. The size consistency is a property that guarantees the consistency of the energy behaviour when interaction between the involved molecular system is nullified (for example, by distance). Size consistency in the CC theory, unlike other theories, does not depend on the size consistency of the reference wave function. For example, it can be shown that the restricted HF model is not always size-consistent.

where the expansion coefficients t_i^r are called the amplitudes. The coupled cluster wave function is defined as follows:

$$|\Psi_{CC}\rangle = e^{\hat{T}} |\Psi_0\rangle \,, \tag{1.71}$$

with $e^{\hat{T}} = \hat{I} + \hat{T} + \frac{1}{2}\hat{T}^2 + ...$. The truncation of the excitation operator $\hat{T}$ defines the coupled cluster method: at the second simplest level,[9] we obtain the coupled-cluster singles-and-doubles (CCSD) model, omitting from the cluster operator all terms that involve higher than single and double excitations: $\hat{T} = \hat{T}_1 + \hat{T}_2$. Inserting the expression in 1.71 into the Schrödinger equation, we obtain:

$$\hat{H}e^{\hat{T}} |\Psi_0\rangle = Ee^{\hat{T}} |\Psi_0\rangle \,, \tag{1.72}$$

$$e^{-\hat{T}}\hat{H}e^{\hat{T}} |\Psi_0\rangle = E |\Psi_0\rangle \,. \tag{1.73}$$

Equations for the amplitudes can be obtained from Eq. 1.73 by multiplying by the reference $\langle\Psi_0|$ and excited state determinants:

$$\langle\Psi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = E, \tag{1.74}$$

$$\langle\Psi_i^r|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = 0, \tag{1.75}$$

$$\langle\Psi_{ij}^{rs}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = 0, \tag{1.76}$$

where the left hand side is the expectation value of the so-called similarity transformed Hamiltonian $e^{-\hat{T}}\hat{H}e^{\hat{T}}$ which is non-Hermitian.

Using the Baker-Campbell-Hausdorff formula we can rewrite the equation 1.74 as follows:

$$E = \langle\Psi_0|\hat{H}|\Psi_0\rangle + \langle\Psi_0|[\hat{H},\hat{T}]|\Psi_0\rangle + \frac{1}{2!}\langle\Psi_0|[[\hat{H},\hat{T}],\hat{T}]|\Psi_0\rangle + ... \tag{1.77}$$

This simplifies considerably, since $\langle\Psi_0|\,\hat{T} = 0$, $\langle\Psi_0|\hat{H}\hat{T}_1|\Psi_0\rangle = 0$ due to the Brillouin theorem and $\langle\Psi_0|\hat{H}\hat{T}_i|\Psi_0\rangle = 0$ for $i > 2$ due to the nature of the electronic Hamiltonian. Then we get an expression:

$$E = E_{\mathrm{HF}} + \langle\Psi_0|[\hat{H},\hat{T}_2]|\Psi_0\rangle + \frac{1}{2}\langle\Psi_0|[[\hat{H},\hat{T}_1],\hat{T}_1]|\Psi_0\rangle \,. \tag{1.78}$$

Alternatively, we can find the energy by projecting the CC Schrödinger equation 1.72 onto the reference wave function:

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E\left\langle\Psi_0\middle|e^{\hat{T}}\Psi_0\right\rangle, \tag{1.79}$$

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E\left\langle\Psi_0\middle|(\hat{I} + \hat{T}_1 + ...)\Psi_0\right\rangle, \tag{1.80}$$

$$\langle\Psi_0|\hat{H}e^{\hat{T}}|\Psi_0\rangle = E, \tag{1.81}$$

which is in accordance with the equation 1.74 since $\langle\Psi_0|\,e^{-\hat{T}} = \langle\Psi_0|$. Then we can simplify this expression using the fact that the electronic Hamiltonian contains

[9] The first being the CCD with $\hat{T} = \hat{T}_2$ because CCS does not contribute due to the Brilloiun theorem.

only one- and two-electron operators and the Brillouin theorem:

$$E = \langle\Psi_0|\hat{H}(\hat{I} + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2)|\Psi_0\rangle\,, \tag{1.82}$$

$$E = E_{\mathrm{HF}} + \sum_{i<j}^{N_{occ}} \sum_{r<s}^{N_{virt}} (t_{ij}^{rs} + t_i^r t_j^s - t_i^s t_j^r)\,\langle\Psi_0|\hat{H}|\Psi_{ij}^{rs}\rangle\,, \tag{1.83}$$

$$E = E_{\mathrm{HF}} + \sum_{i<j}^{N_{occ}} \sum_{r<s}^{N_{virt}} (t_{ij}^{rs} + t_i^r t_j^s - t_i^s t_j^r)(\langle ij|rs\rangle - \langle ij|sr\rangle). \tag{1.84}$$

Including $\hat{T}_3$ in the $\hat{T}$ operator leads to the CCSDT method which, as can be deduced from the table 1.1, is too demanding computationally for all but the smallest systems. Alternatively, the triples contribution may be evaluated by perturbation theory and added to the CCSD results. This leads to the method with the acronym CCSD(T). In this method, the triples contribution is calculated from the formula given by MP4, but using the CCSD amplitudes instead of the perturbation coefficients for the wave function corrections and adding a term arising from fifth-order perturbation theory, describing the coupling between singles and triples. This is the standard scheme for getting very accurate energies.

Let us finish this chapter by noting that the steep formal scaling of the discussed methods as summarized in Tab 1.1 arises from the use of spatially delocalized canonical one-electron wave functions. These are convenient because they are mutually orthogonal and diagonalize the Fock matrix. However, their use does not allow to exploit the fact that electronic correlation (in nonmetallic systems) is an intrinsically short-ranged phenomenon. The use of spatially localized orbitals allows for the construction of MP2 and CC algorithms that scale more favorably with system size, even down to a linear scaling, at the price of a significant increase in complexity with respect to their canonical counterparts [27, 28].

Table 1.1: Formal scaling behavior of the discussed quantum chemical methods (N denotes the number of basis functions) [26].

Method	Scaling
HF	N^4
MP2	N^5
MP3, CCSD	N^6
MP4, CCSD(T)	N^7
MP5, CCSDT	N^8

2. Computational techniques

In this chapter, we describe the fragment and the PBC schemes along with the settings used for the calculations. We also compare the cluster and PBC implementations for the calculation of dimer interaction energies.

2.1 Fragment approach

This approach is based on the idea of decomposing a calculation on a large (theoretically infinite) system into many smaller subsystem calculations. After picking a central molecule labeled as "1", the binding (or lattice) energy of a crystal is then expressed in terms of 2-body, 3-body, 4-body, etc., contributions:

$$E_{\mathrm{b}} = \frac{1}{2}\sum_{1\neq i} E_{1i} + \frac{1}{3}\sum_{1\neq j\neq k} E_{1jk} + \dots . \tag{2.1}$$

The sums involve dimers, trimers, tetramers, etc., within a specified distance from the central molecule based on the speed of convergence and the requested precision of a converged resulting value. It readily follows that this approach can reduce computational costs substantially, if the convergence is fast with respect to the number of molecules in the fragment and with the intermolecular distance.

For practical calculations, we used the Molpro program which uses gaussian-type basis sets to represent orbitals. Specifically, we used the Dunning correlation consistent basis sets which are designed for converging post-Hartree-Fock calculations such as MP2 and CCSD(T) systematically to the complete basis set (CBS) limit [29]. Most calculations have been performed using the aug-cc-pVTZ [1] (for MP2 and CCSD(T)) basis set although aug-cc-pVQZ was also used for some HF calculations, and aug-cc-pVDZ was used for very demanding CCSD(T) calculations of larger clusters. The "aug" basis sets contain diffuse functions which are important for correlated calculations. We demonstrate the superiority of aug-cc-pVXZ (AVXZ) basis sets over cc-pVXZ (VXZ) basis sets for MP2 calculations in Fig. 2.1.

The gaussian-type basis functions efficiently represent molecular orbitals in practice. However, the calculation of dimer interaction energies presents several obstacles. First, the usage of non-orthogonal basis sets leads to the basis set superposition error (BSSE). The BSSE is caused by the fact that when the energies of monomers are evaluated, a smaller basis set is used compared to the basis set used in the dimer energy calculation. The use of a smaller basis set increases the energy of the monomers and leads to overestimated binding. To alleviate this issue, the counterpoise correction can be used, in which the energies of the monomers are evaluated in the complete basis set of the dimer. This means that all calculations are performed using the same basis set and, for monomers, with the help of "ghost orbitals", *i.e.*, basis set functions with zero charge.

The convergence of counterpoise corrected Hartree-Fock energies is shown in Fig. 2.2. One can see that even with the CP correction, there is still a finite deviation from the CBS limit for each employed basis set even though the deviations

[1] The cc-pVXZ label stands for the Dunning correlation-consistent, polarized, valence-only, X-zeta basis [29].

decrease with the basis set size. This is called the basis set incompleteness error (BSIE), as it corresponds to the use of a finite basis set. The BSIE is especially prominent for correlated calculations such as MP2 and CCSD(T). This is caused by the necessity to model the two-electron cusp, a singularity in the derivative of the electronic wave function, which causes the correlation energy to converge as L^{-3} with the largest angular momentum L contained in the basis set.

The convergence of MP2 and CCSD(T) with respect to the basis size can be accelerated by using the explicitly correlated F12 method [30, 31]. This method explicitly introduces a dependence on the interelectron distance into the wave function and allows us to minimize the convergence problems associated with the aforementioned "interelectron cusp" [32]. The comparison of MP2 and MP2-F12 convergence is depicted in Fig. 2.1 for a dimer of methanol. In this figure, one can notice the apparently slower convergence of MP2 compared to MP2-F12. Therefore, when one uses the canonical MP2 scheme, *i.e.*, without F12, an extrapolation of MP2 energies to the complete basis set (CBS) limit is needed to obtain an accurate MP2 interaction energy. Since the total interaction energy is constructed as a sum of the Hartree–Fock and correlation interaction energies, the latter of which converges more slowly with the basis set size, the extrapolation to the CBS limit can be done separately for both components. We employ the frequently used scheme of Helgaker et al. [33]:

$$E_{\mathrm{X}}^{\mathrm{HF}} = E_{\mathrm{CBS}}^{\mathrm{HF}} + Ae^{-BX}, \tag{2.2}$$

$$E_{\mathrm{X}}^{\mathrm{MP2}} = E_{\mathrm{CBS}}^{\mathrm{MP2}} + CX^{-3}, \tag{2.3}$$

with X denoting the cardinal number of a given basis set (D=2, T=3, etc.). The Fig. 2.1 shows that for MP2-F12, the deviation from the CBS limit obtained by an extrapolation is only 0.04 kJ/mol (less than 0.2 %) using the AVTZ basis set and 0.2 kJ/mol (cca 0.8 %) using the AVDZ basis set. Therefore, in this thesis, we employ MP2-F12 and expect accurate results that are very close to the usual MP2 CBS extrapolations. Moreover, concerning molecular solids, we expect results with uncertainties emanating from the BSIE which are lower than 1 %.

Similarly, the superiority of CCSD(T)-F12 convergence over plain CCSD(T) is manifested in Fig. 2.3 for a dimer of methane. However, there is no direct F12 correction to the perturbative triples correction of CCSD(T)-F12 implemented in Molpro, and therefore the basis set incompleteness error of the triples is not improved by the F12 method. A simple improvement of the triples energy is obtained by scaling the triples energy contribution [29]. This is also showed in the aforementioned Figure 2.3. Taking the CBS limit as the result of CCSD(T)-F12 with triples scaling at the AV5Z level, the deviation of AVTZ and AVDZ results from the CBS limit is less than 0.02 kJ/mol (cca 0.7 %). This provides a basis for our error estimates concerning calculations on molecular solids.

For both HF and MP2 calculations is Molpro, density fitting was used to speed up calculations.[2]

[2]Using the two-electron integral notation in Eq. 1.68, we can define electron densities $\rho_{ij}(\vec{r})$ so that

$$\langle ab|rs\rangle = \int \rho_{ar}(\vec{r}_1)\frac{1}{|\vec{r}_1 - \vec{r}_2|}\rho_{bs}(\vec{r}_1)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2. \tag{2.4}$$

The densities may be approximated using an auxiliary basis set as: $\bar{\rho}_{ar} = \sum_P^{N_{fit}} d_P^{ar}\varphi_P(\vec{r})$, where d_P^{ar} are the fitting coefficients.

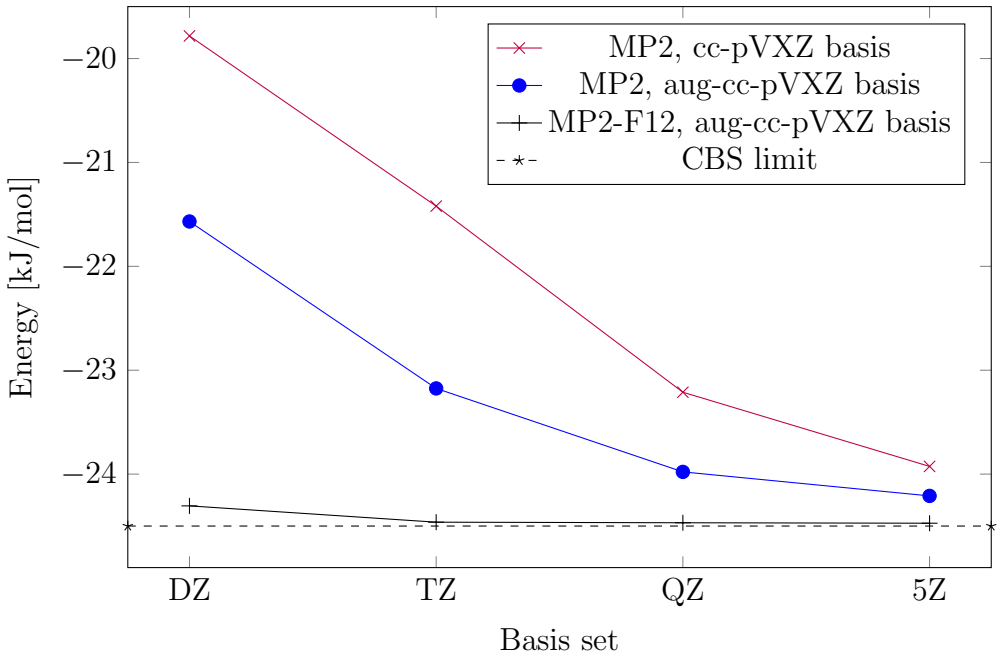

Figure 2.1: Convergence behavior of the total MP2 corrected interaction energy of the methanol dimer for standard Dunning correlation consistent (cc) basis sets, basis sets with added diffuse functions (aug) and with the F12 corrections. See text for the explanation of the CBS limit calculation.

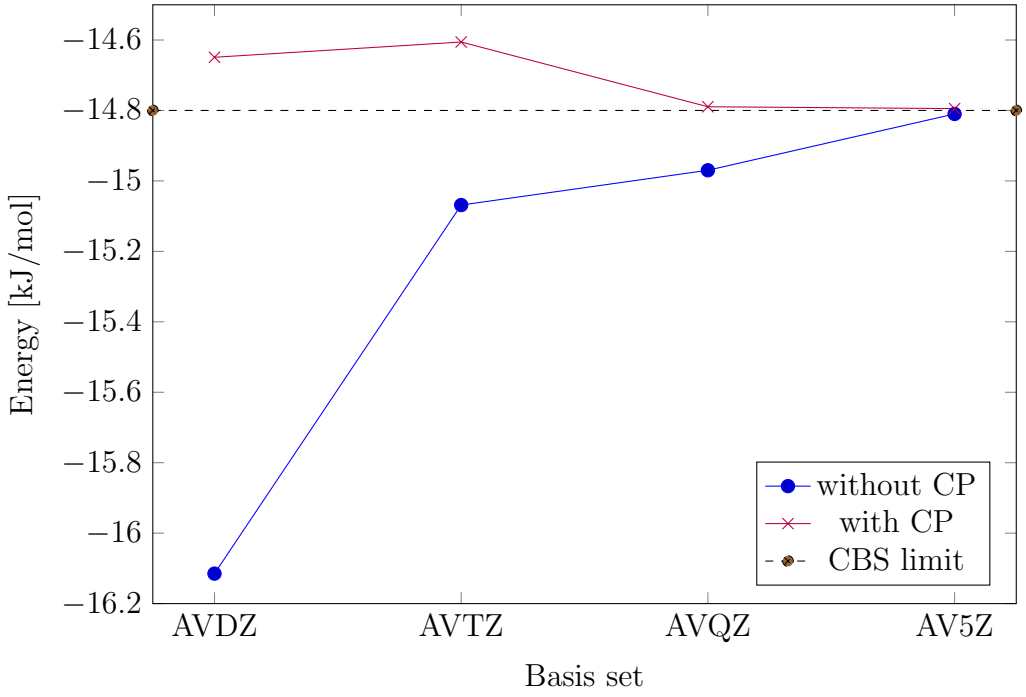

Figure 2.2: Convergence behavior of the Hartree-Fock interaction energy of MeOH dimer with and without the counterpoise correction using the aug-cc-pVXZ basis sets.

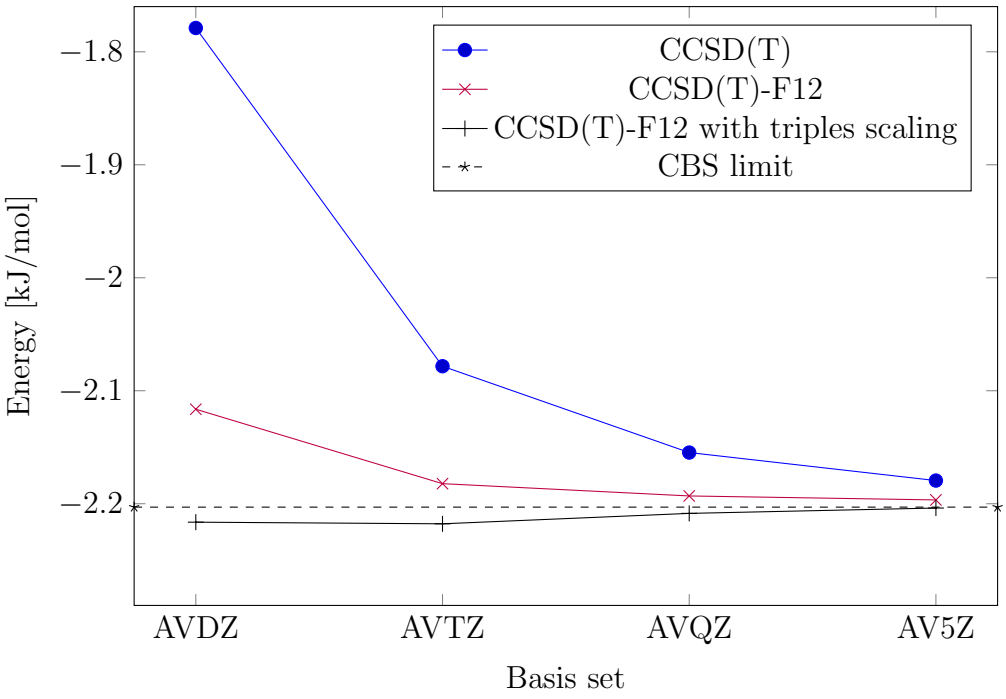

Figure 2.3: Convergence behavior of the CP corrected CCSD(T) interaction energy of the Me dimer. CBS limit is taken as CCSD(T)-F12 with triples scaling using the AV5Z basis set.

2.2 Periodic boundary conditions approach

Periodic boundary conditions (PBC) ought to be more natural for studying periodic systems compared to the finite cluster calculations, especially for systems with low symmetry and large super-cells. To study this hypothesis, we used the *Vienna Ab-initio Simulation Package* [34, 35, 36] to calculate the HF energy and the MP2 correlation energy of our molecular solids. We employed the MP2 implementation by Marsman et al. [37].

In VASP, central quantities, such as the one-electron orbitals, are expressed in plane wave basis sets. Thus, the formalism developed in the Chapter 1.1.4. applies. To recapitulate, integrals in real space over a large system are replaced by integrals over the first Brillouin zone in reciprocal space and these integrals are then approximated by finite sums over the k-points grid in the first Brillouin zone. The interactions between the electrons and ions are described using the projector-augmented-wave method[38, 39]. To determine the electronic groundstate, VASP makes use of efficient iterative matrix diagonalisation techniques, like the residual minimisation method with direct inversion of the iterative subspace (RMM-DIIS) or blocked Davidson algorithms. These are coupled to highly efficient Broyden and Pulay density mixing schemes to speed up the self-consistency cycle [36].

In the PBC approach, the binding energy of a crystal is given by the expression:

$$E_{\mathrm{b}} = E_{\mathrm{crys}}/Z - E_{\mathrm{mol}}, \tag{2.5}$$

where E_{crys} is the energy of a crystal, Z is the number of molecules in the unit cell and E_{mol} is the energy of an isolated molecule. In VASP, the calculated energy of

a crystal depends on the number of k-points and the maximum (cutoff) energy of the plane-wave basis set. Moreover, the energy of the isolated molecule depends on the cell volume used for the calculation. Therefore, to obtain a converged lattice energy it is necessary to extrapolate with respect to these parameters. Our general strategy is to obtain volume or k-point converged energies for a set of cutoffs and then extrapolate to infinite cutoff. For the HF energy we assumed the convergence behavior in the form $1/V$ with the volume of the cell V for an isolated molecule and $1/N_k$ with the number of k-points [40]. This convergence can be accelerated using the Coulomb cutoff technique (tag HFRCUT=-1 in VASP [41]), which we decided to employ. The convergence of the HF energies with respect to the basis set cutoff is rather fast and requires only a sufficiently high cutoff to be used.

For the MP2 energy we also relied on extrapolations to infinite cell volume. For solids, no clear convergence behavior was identified and we simply used a high enough k-point set to assure converged data. The MP2 energies should converge as $\mathrm{ENCUT}^{-\frac{3}{2}}$, but this behavior is not always visible when the energies of solids and molecules are subtracted to form the lattice energy.

The memory requirements of the MP2 calculations grow significantly with the basis size and the number of k-points or the cell volume. This means that the converged energy of the solid $E_\infty^{\mathrm{large}}$ for a large cutoff cannot be obtained directly. To resolve this issue, we first obtained a k-point grid converged energy $E_\infty^{\mathrm{small}}$ for a small basis-set size. This value was corrected for the finite basis-set size using two calculations with a small number of k-points, one with the small basis set E_0^{small} and one with a large basis set E_0^{large} as follows [10]:

$$E_\infty^{\mathrm{large}} = E_\infty^{\mathrm{small}} + (E_0^{\mathrm{large}} - E_0^{\mathrm{small}}). \tag{2.6}$$

For systems with higher numbers of electrons including MeOH and CO_2, it is impossible to calculate MP2 energies in VASP using the standard algorithm which has a quintic (N^5) scaling [42]. In fact, using more than 10 k-points the random access memory (RAM) requirements reach tens of TB. Therefore, converged MP2 data were obtained as a sum of two contributions exhibiting a different convergence behavior: an exchange contribution (xMP2) and a Hartree contribution (dMP2). The xMP2 values converge more rapidly with the energy cutoff and the number of k-points compared to dMP2. To compute dMP2 contributions we employed the random-phase approximation (RPA) algorithm in VASP [43], which scales as only N^3 with the system size and has lower memory requirements compared to the conventional MP2 algorithm which exhibits N^5 scaling.

2.3 Dimer benchmarks

To provide a foundation for our lattice energies comparison, we first seeked to obtain matching values of interaction energies of selected benchmark dimers in both Molpro and VASP. First, we checked that with our set-up we reproduce the results for both MP2 and CCSD(T) published in Jurečka et al. [44] and Řezáč et al. [45] for methane and methanol dimers in the respective geometry. Using the MP2-F12 method with the AV5Z basis set, we obtained MP2 energies within 0.06 kJ/mol (methane) and 0.02 kJ/mol (methanol) from the reference values

(see Table 2.1). Furthermore, we calculated CCSD(T) interaction energies and compared them to MP2. The CCSD(T)-F12 method with triples scaling was used with AV5Z (for methane) or AVQZ (for methanol) basis sets. Our results agree very well with the reference values (see Table 2.2). We found that CCSD(T) increases the binding energies slightly with respect to MP2. Specifically, we found an increase of 0.15 kJ/mol for methane and 0.06 kJ/mol for methanol. These small corrections indicate that the two-body contribution to the lattice energy should be described rather well by MP2.

Table 2.1: Dimer HF and MP2 interaction energies (in kJ/mol) using the finite cluster and PBC approaches. The reference data are from [44].

	HF		MP2 corr.		HF + MP2		
	clus.	PBC	clus.	PBC	clus.	PBC	Ref.
Me	1.53	1.53	−3.59	−3.50	−2.06	−1.97	−2.13
MOH	−14.79	−14.79	−9.66	−9.51	−24.45	−24.3	−24.43
NH_3	2.05	2.08	−2.07	−2.06	−0.02	0.03	–
CO_2	−0.27	−0.26	−3.87	−3.75	−4.14	−4.01	–

Table 2.2: Dimer CCSD(T) interaction energies (in kJ/mol) using the finite cluster approach. The reference data are from [44].

	CCSD(T)	
	clus.	Ref.
Me	−2.21	−2.22
MeOH	−24.51	−24.48
NH_3	−0.35	–
CO_2	−4.30	–

In VASP, we found that Hartree-Fock energies agree nicely with the results from the cluster approach without any problems with convergence. Concerning MP2, however, we encountered a slow convergence behavior of MP2 correlation energies with respect to the cell volume, see Figures 2.4 and 2.5. For extrapolations of $1/V$ to the infinite cell volume, we used cells with sides up to 14 Å. Larger cells are technically demanding and computationally expensive. Based on the spread of the extrapolated data, as specified in Table 2.3, we estimated the uncertainty of the converged result for both methane (0.03 kJ/mol) and methanol (0.01 kJ/mol). The convergence behavior of the cell-size extrapolated MP2 binding energies with respect to ENCUT and extrapolations of the form $\mathrm{ENCUT}^{-\frac{3}{2}}$ are presented in Figures 2.6 and 2.7 for Me and MeOH, respectively. For methane, we see from the resolution of the graph that the results are already well-converged even for ENCUT = 1000 eV and our final estimate is -3.50 ± 0.03 kJ/mol. For methanol, however, a further extrapolation of the form $\mathrm{ENCUT}^{-\frac{3}{2}}$ seems to be reasonable and the result is -9.51 ± 0.02 kJ/mol. From Table 2.1 we conclude that VASP underestimates correlation effects in both methane and methanol. Compared to the cluster approach, MP2 corrections are lower by 0.1 – 0.2 kJ/mol. This can be caused by inaccuracies of the PAW potentials used in VASP [39]. Another contribution are the uncertainties related to the extrapolation to the infinite

volume. The relative error of 4.5% for the total HF and MP2 energy of methane is rather high due to opposite HF and MP2 contributions to binding.

Table 2.3: The extrapolation of methane and methanol dimer interaction energies (computed with the basis-set energy cutoff 1000 eV) to the infinite cell size as $\frac{1}{V}$ for four different ranges of data.

Cell sizes ($\sqrt[3]{V}$) [Å]	$E_{\text{dim.}}^{\text{extrap.}}$	
	Me	MeOH
12 – 14	−3.470	−9.606
12.5 – 14	−3.488	−9.618
13 – 14	−3.499	−9.613
13.5 – 14	−3.514	−9.617

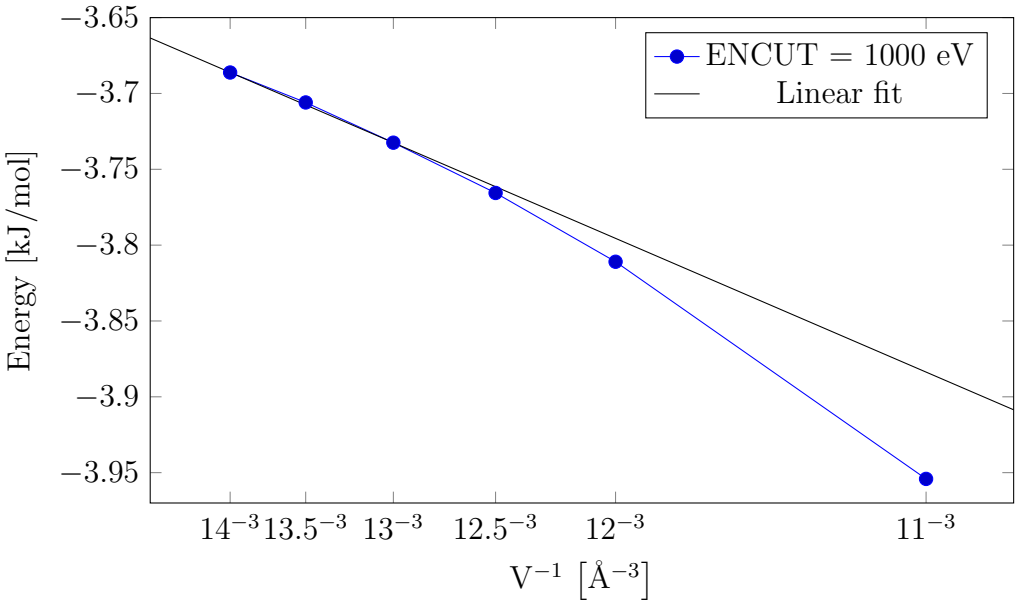

Figure 2.4: Convergence of MP2 binding energies of the Me dimer with respect to the cell size for the basis-set cutoff energy (ENCUT) of 1000 eV. Only three largest cell sizes of 13, 13.5, and 14 Å are used for the regression.

Next, we selected one dimer from both NH_3 and CO_2 solids, and compared the values of interaction energies obtained by both approaches to see how fast both methods converge and what level of precision can be expected in succeeding calculations. From the ammonia solid, we selected a dimer without a hydrogen bond (see Fig. 2.8). Performing the Hartree-Fock calculations, we found a slow convergence with the cell size quite analogously to methane and methanol dimers. The difference between the PBC approach and the cluster scheme is 0.03 kJ/mol or 2 % which is acceptable for our purposes. We found no binding within the MP2 correlation scheme using both approaches; however, CCSD(T)-F12 calculations showed a non-negligible binding contribution, which indicates the errors one can expect from the less accurate MP2 method. The resulting data can be found in Table 2.1. For the selected dimer of CO_2, we found that HF energies agreed well

for both approaches, but MP2 in VASP understimated the correlation energy by 0.13 kJ/mol, which is about 3% of the MP2 contribution. The possible reasons for this discrepancy are again less accurate PAW potentials or an inaccurate extrapolation to the infinite cell volume.

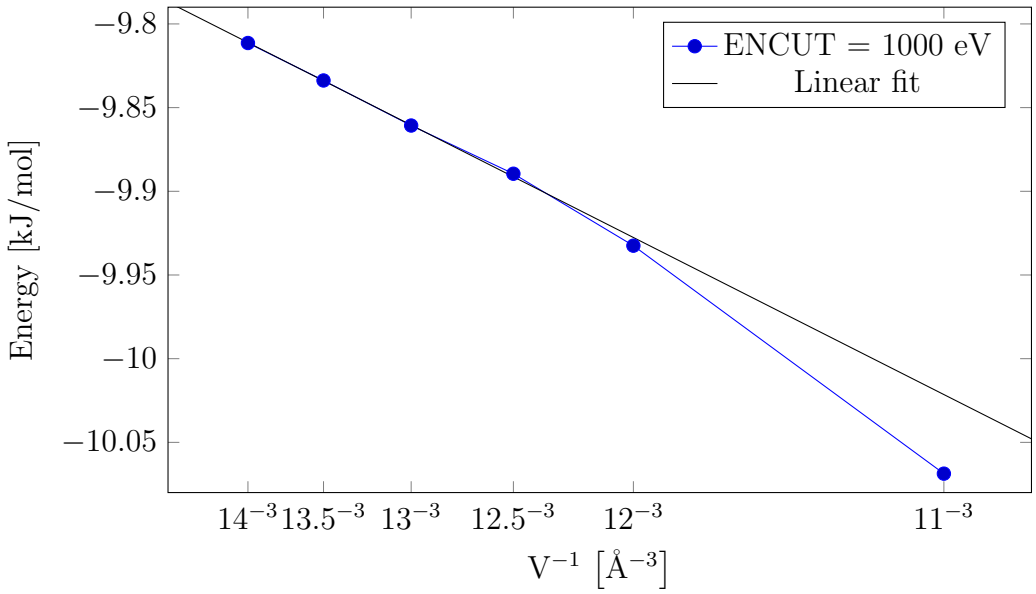

Figure 2.5: Convergence of MP2 binding energies of the MeOH dimer with respect to the cell size using ENCUT = 1000 eV. Only three largest cell sizes of 13, 13.5, and 14 Å are used for the regression.

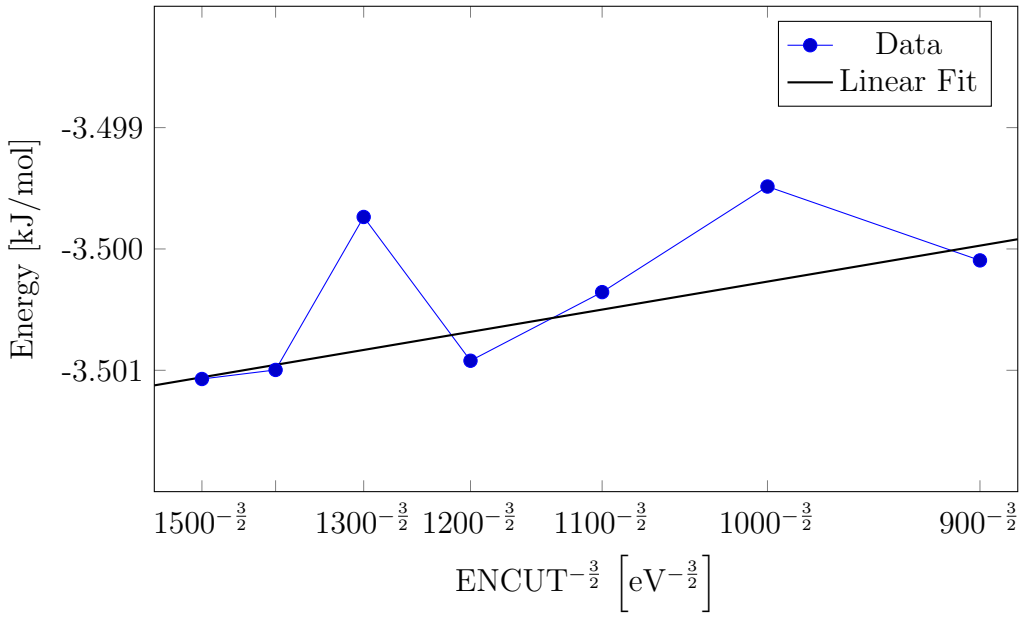

Figure 2.6: Convergence behavior of the cell-size extrapolated MP2 binding energies of Me dimer.

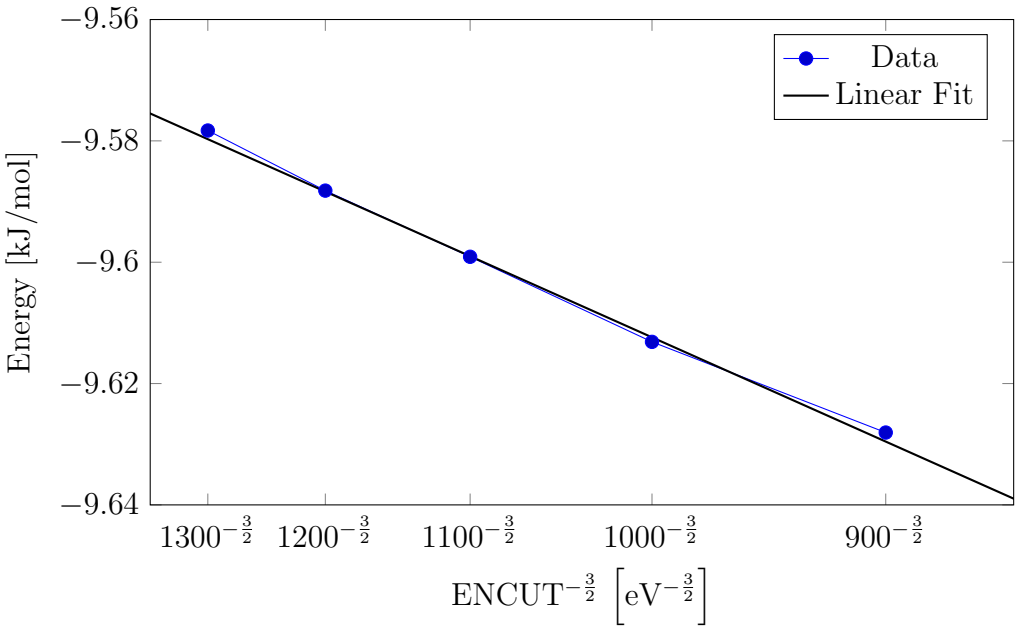

Figure 2.7: Convergence behavior of the cell-size extrapolated MP2 binding energies of MeOH dimer.

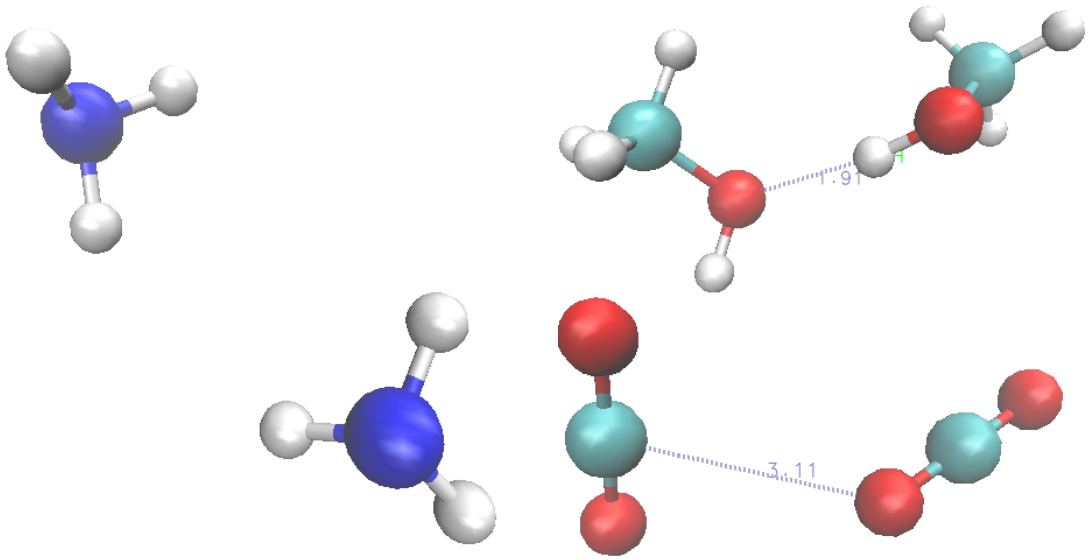

Figure 2.8: The geometries of dimers of ammonia (left), methanol (upper right) and CO_2 (lower right) with indicated bond lengths (in Å).

3. Results for molecular solids

In this chapter we present the results obtained for lattice energies with the fragment and PBC schemes. At the end, we compare our results with the experimental values and the results from literature.

The studied systems have the following properties.[1] Methane (phase I), an fcc crystal with lattice constant $a = 5.84\,\text{Å}$ [9], has 1 molecule in the unit cell, and zero dipole moment. Methanol ($\alpha-$phase), an orthorhombic crystal with lattice constants $a = 4.873\,\text{Å}$, $b = 4.461\,\text{Å}$, $c = 8.867\,\text{Å}$ [47], has 4 molecules in the unit cell, and dipole moment of 1.69 D. Carbon dioxide (phase I), an fcc crystal with lattice constant $a = 5.624\,\text{Å}$ [48], has 4 molecules in the unit cell, and zero dipole moment. Ammonia (phase I), an fcc crystal with lattice constant $a = 5.13\,\text{Å}$ [49], has 4 molecules in the unit cell, and dipole moment of 1.46 D.

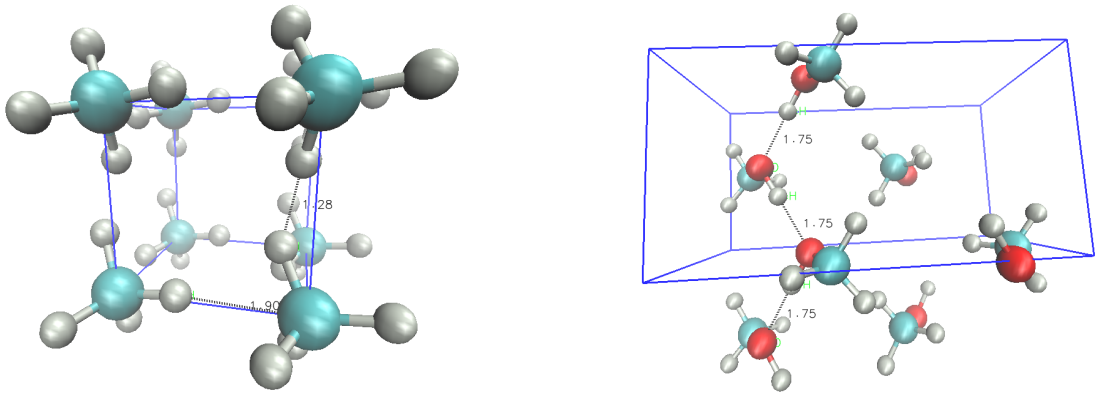

Figure 3.1: The structures of crystalline methane (left) and methanol (right) with several indicated bond lengths in Å.

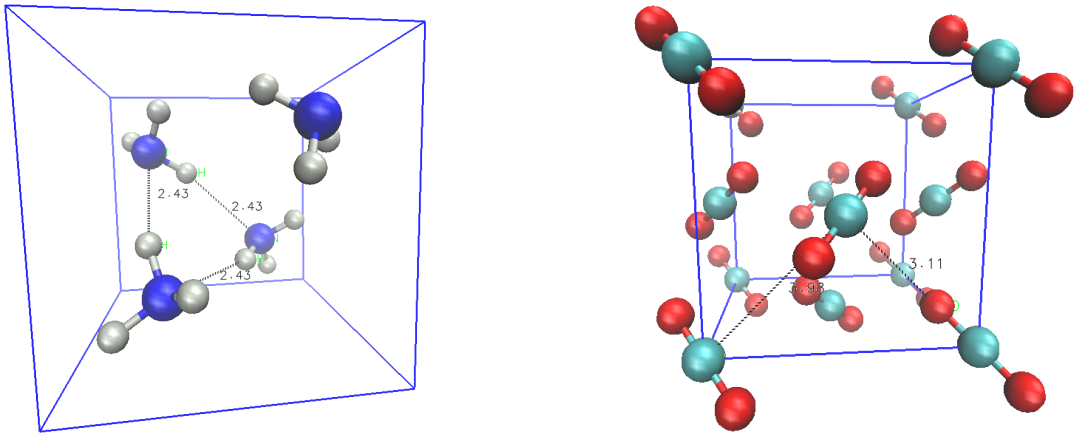

Figure 3.2: The structures of crystalline ammonia (left) and carbon dioxide (right) with several indicated bond lengths in Å.

[1]The geometry data were obtained from the Crystallography Open Database `www.crystallography.net/cod/` [46].

3.1 Methane

Fragment approach

Thanks to the non-polar nature of the methane molecule, the fragment expansion as defined in Eq. 2.1 converged rather quickly. We decided to truncate the fragment expansion with tetramers, which we expected to contribute marginally. We found that pairwise interactions constitute more than 95% of the total lattice energy using any of our methods. Moreover, this contribution converges rather quickly with distance in accordance with the theoretical prediction of a dispersion force dying off as $\frac{1}{r^6}$. This is demonstrated in Fig. 3.3.

Whereas trimers and tetramers contribute less than 1% to the total MP2 energy, CCSD(T) calculations involving perturbative triples increase this share to roughly 5%. This can be simply understood from the fact that within the MP2 scheme only two electrons can be simultaneously excited. The results of our calculations can be found in Tables 3.1 and 3.2.

Taking into account dimers up to 30 Å from the central molecule, trimers up to 10 Å, and closest[2] tetramers, we found the HF lattice energy of Me to be equal to $5.06 \pm 0.05\,\mathrm{kJ/mol}$. This energy is positive, *i.e.*, the crystal is not bound in the HF approximation and binding arises from electron correlations that cause van der Waals dispersion. Adding MP2 and CCSD(T) corrections, we found the total lattice energy of $-9.8 \pm 0.2\,\mathrm{kJ/mol}$ and $-10.5 \pm 0.2\,\mathrm{kJ/mol}$, respectively.

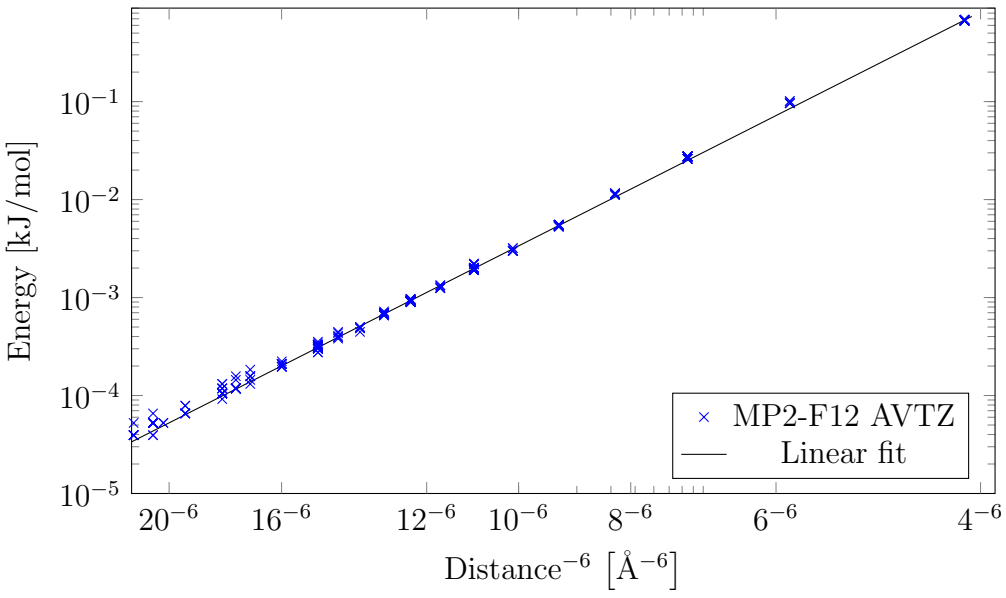

Figure 3.3: Dimer interaction energy decreases as distance to the sixth. Here, MP2-F12 data using the AVTZ basis set are presented for dimers in solid Me.

The stated errors are based on our tests on dimers, which are described in more detail in the previous chapter. For solids, we used explicitly correlated F12 implementations and the aug-cc-pVTZ (AVTZ) basis set, which gives a basis set

[2]Here, we picked 3 strongest trimers and to each trimer we added ca 20 closest molecules up to 5 Å away to form tetramers. Seeing that this contributed so little to the total energy, compared with our error margin, we did not consider any further fragments.

incompleteness error estimate of 2% (see Fig. 2.1) for MP2 and 1% for CCSD(T). For computionally demanding CCSD(T) calculations on trimers and tetramers, we used the AVDZ basis set to obtain results converged with respect to the cutoff distance. We repeated the calculations on 10 trimers using the AVTZ basis set and we found a deviation of 2% from the AVDZ results. In fact, this changes the Δ[CC-MP2] contribution from trimers by only 0.01 kJ/mol, which is negligible compared to our total error margin.

PBC approach

The HF energies of the Me solid converged rather quickly with the number of k-points and the cell volume. Concerning the convergence with respect to the basis set size, we found that a precision of 0.1 kJ/mol is apparently reached with the basis set cutoff (ENCUT) value of 1500 eV, as shown in Fig. 3.4. The resulting HF lattice energy was found to be equal to 5.1 ± 0.1 kJ/mol in accordance with the fragment approach.

Performing MP2 calculations, we confronted a rather slow and erratic convergence behavior with respect to the basis set size (Fig. 3.5). However, note the scale of the graph. The results are essentially well-converged already for the smallest cutoff of 1000 eV. Thanks to the small number of electrons in the unit cell, we were able to perform MP2 calculations using the standard MP2 algorithm even for a large grid of $4 \times 4 \times 4$ k-points (with ENCUT = 700 eV). Using the extrapolation techniques discussed in Chapter 2, we obtained the value of the MP2 correction to the lattice energy of Me -15.1 ± 0.1 kJ/mol in a good agreement with the fragment result. The uncertainty value is our estimate based on the speed of convergence as shown in Fig. 3.5.

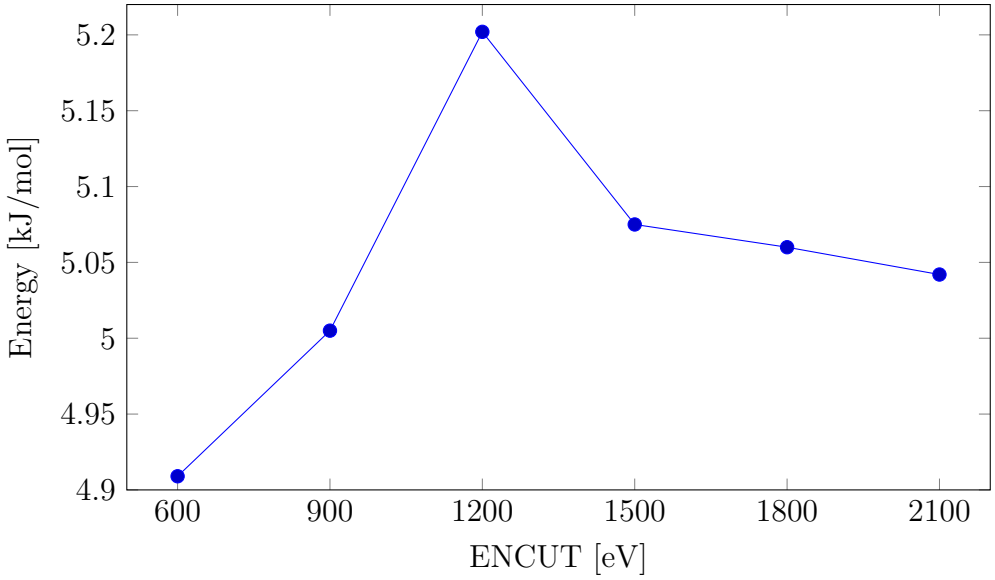

Figure 3.4: Convergence behavior of the cell-size and k-point extrapolated HF binding energies of solid Me.

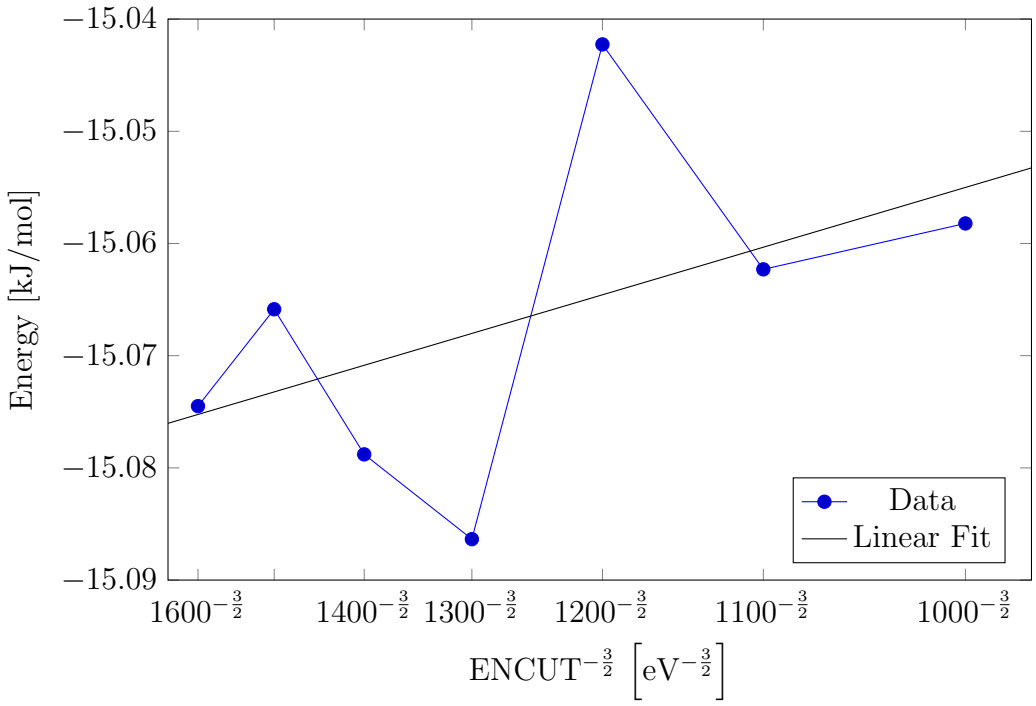

Figure 3.5: Convergence behavior of the cell-size and k-point extrapolated MP2 binding energies of solid Me. The extrapolated value is −15.10 kJ/mol.

3.2 Methanol

Fragment approach

During the analysis of the MeOH crystal geometry using the VMD program, we discovered hydrogen-bonded parallel zig-zag chains along the y-axis. Dimers, trimers, and marginally tetramers along such a chain contribute considerably (more than 60%) to the total HF lattice energy. Compared to methane, the many-body expansion of the HF energy for MeOH does not converge so rapidly. In contrast to methane, we found contributions from trimers to constitute more than a third of the HF energy of the solid. Moreover, this contribution diminishes more slowly with distance. Since the number of trimers that one needs to account for quickly increases with the cut-off distance and the cost of computing trimer contributions is 6 times higher compared to dimers, this ultimately leads to a higher uncertainty of the final value. For the total HF lattice energy we obtained the value of $-17.6 \pm 0.4\,\mathrm{kJ/mol}$, which includes dimers within 15 Å, trimers within 10 Å from the central molecule and 5 tetramers constructed from molecules exhibiting the strongest mutual interaction.

The MP2 correction using the AVTZ basis set mostly affects the contribution from dimers. We also discovered a small (less than 1 kJ/mol) MP2 correction from trimers. This comes mostly from short-range hydrogen bonds because we did not include trimers further than 10 Å from the central molecule. Two examples of strong hydrogen-bonding with opposite effects are shown in Fig. 3.6. Summing up the trimer contribution, we found that the anti-binding (positive energy) trimers dominate. The role of tetramers proved to be negligible (a few hundredths of

kilojoule per mole). For exact results, see Tab. 3.1. These corrections originating from hydrogen bonding go to zero with increasing distance of molecules very quickly. Our final estimate of the MP2 contribution to the lattice energy is $-37.4 \pm 0.8\,\mathrm{kJ/mol}$ with cca 97% coming from pairwise interactions. The error estimate is based on our dimer tests, as we discussed in the methane section. At the same time, we allow for a small error margin coming from considering a finite number of trimers.

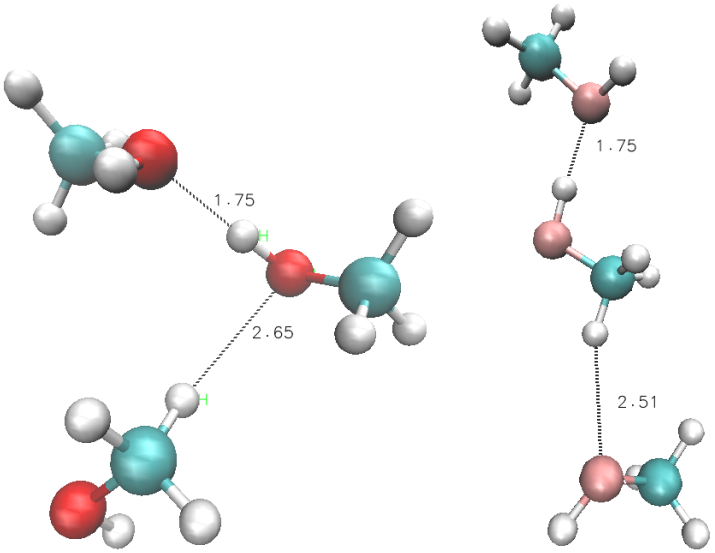

Figure 3.6: Two examples of methanol trimers with opposite binding effects to the lattice energy. The trimer on the left contributes −0.47 kJ/mol, −0.45 kJ/mol, and −0.37 kJ/mol at the HF, MP2, and CCSD(T) levels of theory, while the trimer on the right supplies 0.55 kJ/mol, 0.50 kJ/mol, and 0.47 kJ/mol to the total HF, MP2, and CCSD(T) lattice energy. The calculations were done using the F12 method with the AVTZ basis set.

Adding the CCSD(T) correction to MP2 for dimers using the AVTZ basis set we found the total lattice energy of $-55.6 \pm 0.8\,\mathrm{kJ/mol}$. However, CCSD(T) calculations on trimers were found to be rather expensive even with the AVDZ basis set (1 trimer takes more than an hour on 6 CPUs at Metacentrum), therefore, we only included cca 300 of them. Morevoer, we only performed a few CCSD(T) calculations on tetramers due to a high computational cost even with the AVDZ basis set. Still, we checked that these tetramers contribute less than a hundredth of a kilojoule per mole.

PBC approach

The HF energy converged rather quickly compared to methane and we obtained the value of $-18.1 \pm 0.1\,\mathrm{kJ/mol}$, which is still on par with the fragment value considering the error margin.

MP2 calculations using the standard MP2 algorithm in VASP could not be performed for larger than $2 \times 2 \times 1$ k-point grid due to high memory requirements. Therefore, we employed the RPA algorithm to obtain converged direct (Hartree) contributions to MP2 energies as discussed in Chapter 2. Using this approach, we

obtained well-converged data plotted in Fig. 3.7 and the final value of the MP2 correction of $-37.5 \pm 0.2\,\mathrm{kJ/mol}$, which agrees with the result from the fragment approach.

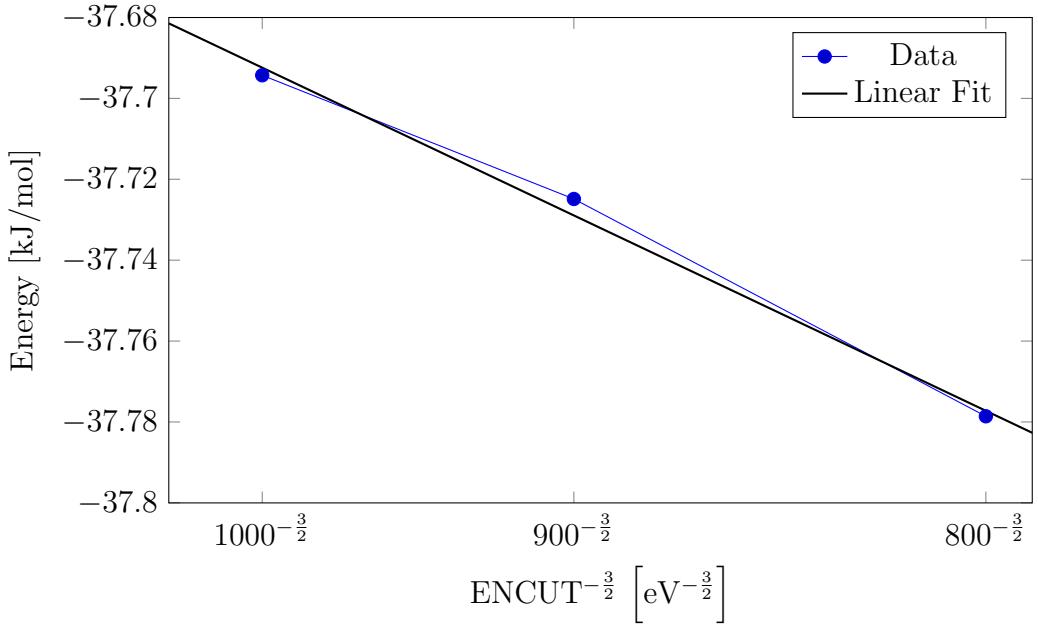

Figure 3.7: Convergence behavior of the cell-size and k-point extrapolated MP2 binding energies of solid MeOH. The extrapolated value is −37.48 kJ/mol.

3.3 Ammonia

Fragment approach

Ammonia is characterized by both weak hydrogen bonds and dispersion interactions. In our calculations, we included dimers within 15 Å from the central molecule since a rather fast convergence with intermolecular distance was observed due to the short-range nature of the dispersion interactions and hydrogen bonds. In contrast to methanol, where we observed a significant (more than 33%) contribution from trimers to the Hartre-Fock lattice energy, in solid ammonia trimers contribute only cca 6% and tetramers less than 0.1% to the HF lattice energy. With MP2 corrections, pairwise interactions dominate even more notably: they constitute more than 99% of the total MP2 corrected lattice energy which we estimate to be equal to $-35.1 \pm 0.4\,\mathrm{kJ/mol}$. Our calculations include dimers within 15 Å, trimers within 10 Å from the central molecule and a few tens of tetramers constructed from 6 strongest trimers.

Whereas at the MP2 level trimers play only a minor role and their summed contribution is lower than 1 kJ/mol, we observed stronger anti-binding tendencies of trimers with the CCSD(T) correction contributing more than 4 kJ/mol to the lattice energy (see Table 3.2). The effect of a given trimer is determined by its geometry. This is demonstrated in Figures 3.10 and 3.11 for three selected trimers. Adding the CCSD(T) correction to MP2 using the AVTZ basis set for dimers

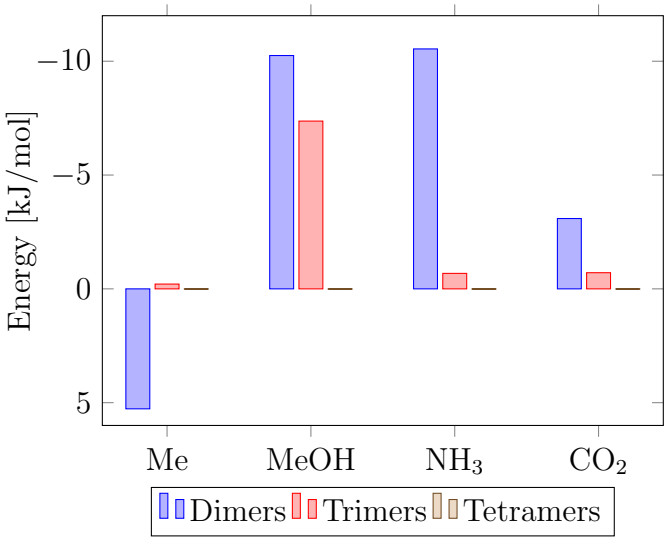

Figure 3.8: Contributions to the Hartree-Fock energy of molecular solids from dimers, trimers and tetramers. Notice the inverted y-axis.

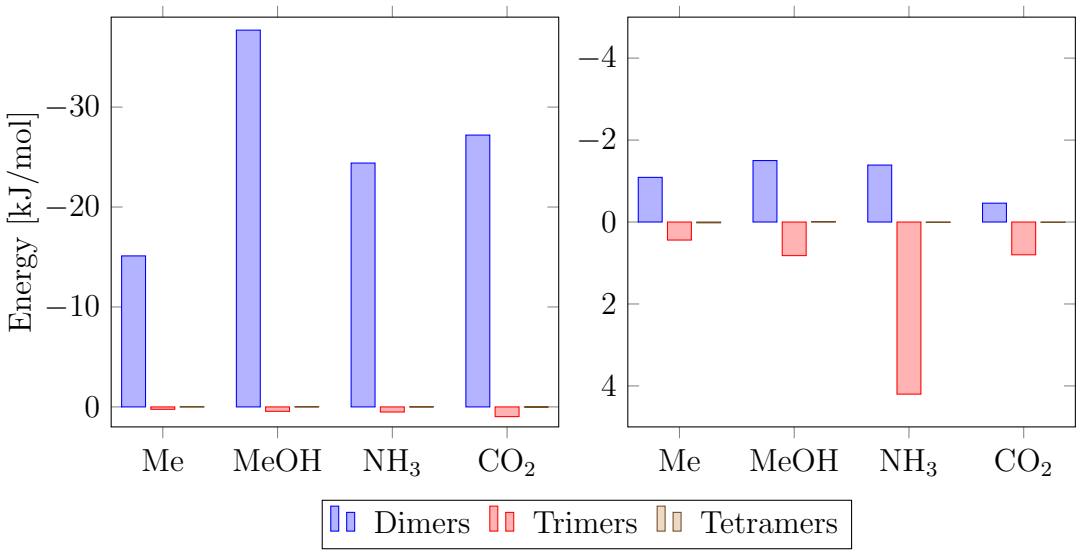

Figure 3.9: Contributions to the MP2 correction (left) and the Δ [CC – MP2] correction (right) to the lattice energies of molecular solids from dimers, trimers and tetramers. Notice the inverted y-axis.

and AVDZ basis set for trimers we found the total lattice energy of $-32.2 \pm 0.5\,\mathrm{kJ/mol}$. The errors represent our estimates of the basis set incompleteness error and we also considered a small error margin emanating from considering a finite number of trimers.

PBC approach

The Hartree-Fock energy converged more slowly with both the cell volume and the k-point grid size for ammonia and we obtained the value of $-11.1 \pm 0.1\,\mathrm{kJ/mol}$ with the error margin estimated from the convergence behavior. This value agrees with the result from the fragment approach ($-11.2 \pm 0.2\,\mathrm{kJ/mol}$).

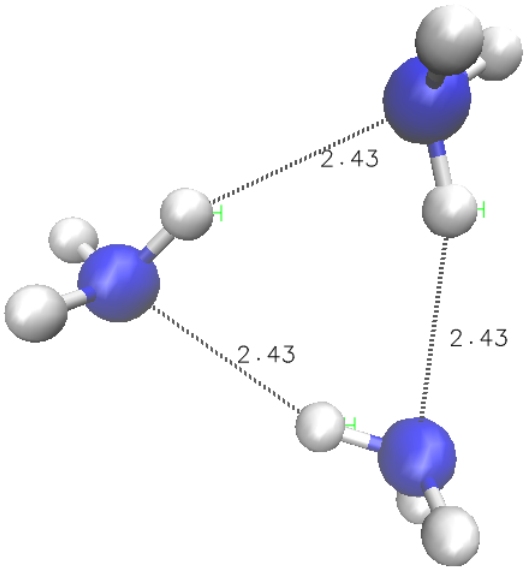

Figure 3.10: An example of a triangle-shaped NH_3 trimer with a strong binding effect due to hydrogen bonds. The trimer contributes -0.78 kJ/mol at the HF and MP2 level and -0.74 kJ/mol at the CCSD(T) level to the lattice energy. The F12 methods and AVTZ basis sets were employed.

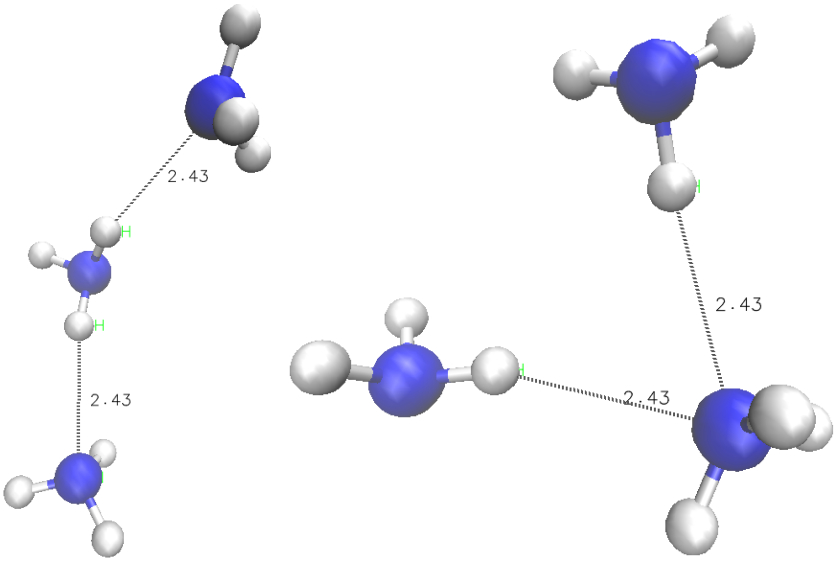

Figure 3.11: Two examples of NH_3 trimers with an anti-binding effect to the lattice energy. The trimer on the left contributes 0.27 kJ/mol at the HF level and 0.25 kJ/mol at both MP2 and CCSD(T) levels of theory, while the trimer on the right supplies 0.11 kJ/mol, 0.15 kJ/mol, and 0.20 kJ/mol to the total HF, MP2, and CCSD(T) lattice energies. The F12 methods and AVTZ basis sets were employed.

MP2 calculations using the standard MP2 algorithm in VASP were performed for grids up to $3 \times 3 \times 3$ k-points (using the largest $3 \times 3 \times 3$ grid the maximum basis set cutoff value was 600 eV) and RPA calculations up to $4 \times 4 \times 4$ with ENCUT = 700 eV. The cell-size and k-point converged data are plotted in Fig.

Table 3.1: HF and MP2-F12 calculations of dimer, trimer and tetramer contributions to total latice energies (in kJ/mol) with AVTZ (AVDZ for tetramers) basis set. See text for error estimates.

	HF			MP2 corr.			HF + MP2		
Solid	dim.	trim.	tetr.	dim.	trim.	tetr.	dim.	trim.	tetr.
Me	5.27	−0.21	0.01	−15.1	0.24	−0.02	−9.84	0.03	−0.01
MOH	−10.2	−7.41	0.01	−37.7	0.44	−0.02	−47.9	−6.97	−0.01
NH_3	−10.5	−0.68	0.02	−24.4	0.51	−0.01	−34.9	−0.17	0.01
CO_2	−3.09	−0.71	0.01	−27.2	0.97	0.00	−30.3	0.26	0.01

Table 3.2: CCSD(T)-F12 fragment calculations (with triples scaling) of interaction energies (in kJ/mol). Dimers calculated with AVTZ basis, trimers and tetramers with AVDZ or AVTZ (Me) basis.

Method	Δ[CC-MP2]			Total CCSD(T)-F12		
Solid	dim.	trim.	tetr.	dim.	trim.	tetr.
Me	−1.09	0.44	0.01	−10.93	0.47	−0.01
MeOH	−1.50	0.82	0.01	−49.44	−6.14	−0.01
NH_3	−1.39	4.29	0.01	−36.3	4.12	0.01
CO_2	−0.46	0.80	0.00	−30.7	1.06	0.01

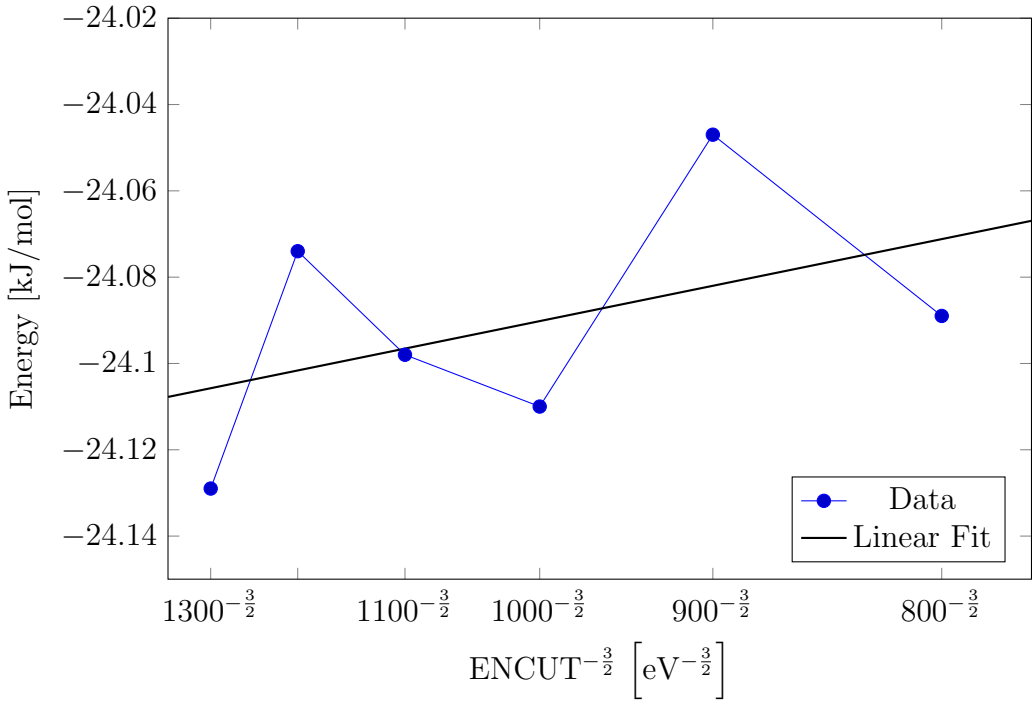

Figure 3.12: Convergence behavior of the cell-size and k-point extrapolated MP2 binding energies of the NH_3 solid. The extrapolated value is −24.14 kJ/mol.

3.12 and the final value of the MP2 correction -24.1 ± 0.2 kJ/mol agrees well with the result from the fragment approach (-23.9 ± 0.4 kJ/mol).

Table 3.3: Calculated values of lattice energies (in kJ/mol) of four molecular solids.

Method	HF		MP2 corr.		HF + MP2		CCSD(T)
Solid	frag.	PBC	frag.	PBC	frag.	PBC	frag.
Me	5.06	5.05	−14.9	−15.1	−9.82	−10.0	−10.5
MeOH	−17.6	−18.1	−37.4	−37.5	−54.9	−55.6	−55.6
NH_3	−11.2	−11.1	−23.9	−24.1	−35.1	−35.2	−32.2
CO_2	−3.80	−3.79	−26.2	−26.3	−30.0	−30.1	−29.7

3.4 Carbon dioxide

Fragment approach

In the fragment expansion we observed a rather fast convergence with trimers and tetramers contributing less than 5% to the total lattice energy. However, at the HF level, trimers contribute more than 20% which is significantly more than 5% in ammonia. We obtained the total HF lattice energy of -3.8 ± 0.1 kJ/mol, which includes dimers within 15 Å, trimers within 10 Å from the central molecule and 5 closest tetramers.

The MP2 correction converged fairly quickly with the cutoff distance and we found the MP2 contribution to the lattice energy to be equal to -26.2 ± 0.4 kJ/mol with more than 95% coming from pairwise interactions. We also included more than 1000 trimers at the AVTZ level, which were found to contribute marginally, *i.e.*, by less than 1 kJ/mol. We also checked 5 closest tetramers with the AVTZ basis set and discovered that their contribution is less than 0.01 kJ/mol and therefore we did not perform further calculations on tetramers.

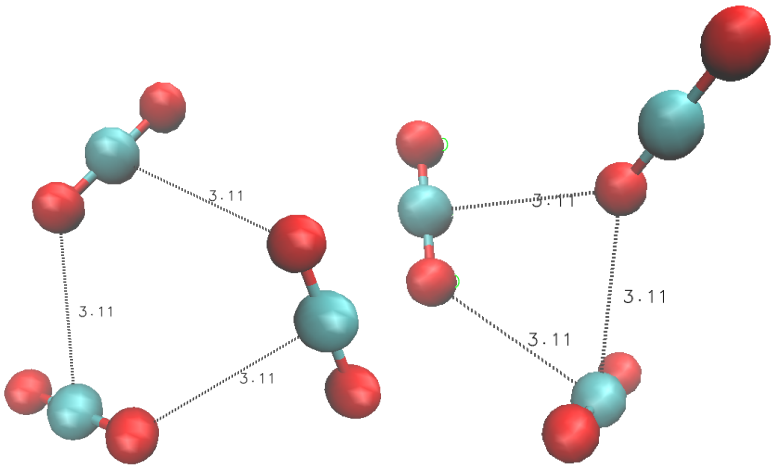

Figure 3.13: Two examples of CO_2 trimers with opposite binding effects to the lattice energy. The trimer on the left contributes −0.1 kJ/mol at both MP2 and CCSD(T) levels of theory, while the trimer on the right supplies 0.04 kJ/mol (MP2) and 0.08 kJ/mol (CCSD(T)) to the lattice energy. The calculations were done using the F12 method with the AVTZ basis set.

Whereas at the MP2 level trimers contribute less than 1 kJ/mol, with the CCSD(T) correction we found a more significant contribution in analogy with

the previous case of ammonia. As we demonstrate in Fig. 3.13, a symmetric (triangular) arrangement of molecules yields a negative (binding) interaction energy, whereas a non-symmetric configuration tends to yield a positive interaction energy. Adding the CCSD(T) correction to MP2 using the AVTZ basis set for dimers and the AVDZ basis set for trimers we found the total lattice energy of $-29.7 \pm 0.5\,\mathrm{kJ/mol}$. Due to a high computational cost of CCSD(T) calculations, we only included cca 50 trimers. Tetramers were too computationally expensive (1 tetramer takes more than 6 hours using 6 CPUs) even using the AVDZ basis set, therefore, we only included 5 tetramers which yielded less than 0.01 kJ/mol as a correction to MP2.

PBC approach

The HF energy converged rather slowly and the use of the Coulomb cutoff technique was appreciated. We estimate an error of cca 0.1 kJ/mol (roughly 3%). Still in agreement with the fragment approach, we obtained the value of $-3.8 \pm 0.1\,\mathrm{kJ/mol}$.

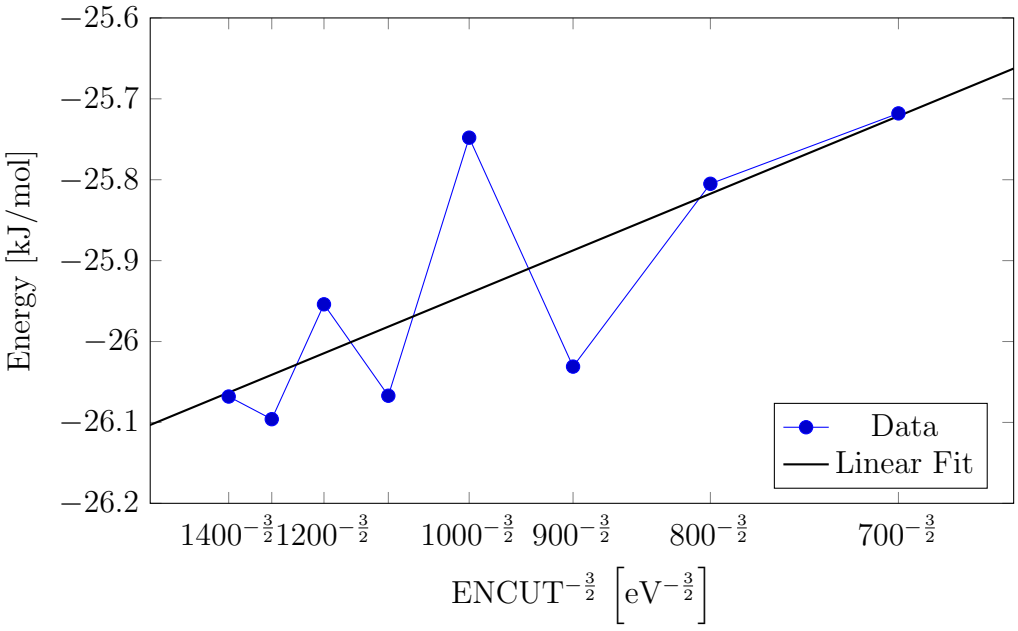

Figure 3.14: Convergence behavior of the cell-size and k-point extrapolated MP2 contribution to the binding energy of CO_2 solid. The extrapolated value is equal to −26.25 kJ/mol.

MP2 calculations using the standard MP2 algorithm in VASP were performed up to $2 \times 2 \times 2$ k-point grid with ENCUT = 800 eV and RPA calculations up to $4 \times 4 \times 4$ with ENCUT = 600 eV. Here, we confronted a slow convergence of MP2 energies with respect to the basis set cutoff value. This is evident from Fig. 3.14. Therefore, the estimated uncertainty is a little higher compared to other solids: $-26.3 \pm 0.3\,\mathrm{kJ/mol}$, which is close to the $-26.2 \pm 0.4\,\mathrm{kJ/mol}$ from the fragment approach.

3.5 Comparison with experimental data

A comparison of our results with previous experimental or theoretical results is not straightforward. In fact, the errors in experimental lattice energies often approach the magnitude of the errors in theoretical calculations [19]. Experimental data need a careful correction for thermal and zero-point effects in order to compare to static HF and post-HF results. Moreover, this correction introduces an uncertainty which, combined with experimental errors, limits the insight that can be gained from the comparison [13]. In Table 3.4 we compare our results with the experiment-based lattice energies taken from the works of Reilly and Tkatchenko [16] (ammonia, carbon dioxide) and Červinka and Fulem [50] (methanol). These were derived from experimental sublimation enthalpies by considering vibrational and thermal contributions. An appropriate value for methane was not found in literature, therefore we also include a comparison with recent fragment-based calculations performed by Červinka et al. [24] at the CCSD(T) level of theory. We see that our CCSD(T) calculated results agree with the reference data derived from experiment and other CCSD(T) results within 1 kJ/mol for MeOH and CO_2. For ammonia, our result underestimates the reference values, which could be caused by different geometry data.

Table 3.4: Calculated values of lattice energies (in kJ/mol) and the reference values derived from experimental sublimation enthalpies by Reilly and Tkatchenko [16] (Ref. A) and Červinka and Fulem [50] (Ref. B). For comparison, recent fragment-based CCSD(T) calculations by Červinka et al. [24] (Ref. C) are also included.

Method	HF + MP2-F12		CCSD(T)-F12	Exp.	Exp.	CCSD(T)
Solid	frag.	PBC	frag.	Ref. A	Ref. B	Ref. C
Me	− 9.82	−10.0	−10.5	–	–	−11.0
MeOH	−54.9	−55.6	−55.6	–	−54.9	−53.8
NH_3	−35.0	−35.2	−32.2	−37.2	−38.5	−39.1
CO_2	−30.0	−30.1	−29.7	−28.4	−28.7	−21.2

Conclusion

In this thesis, we used two different approaches to obtain accurate lattice energies of four molecular solids with different physico-chemical characteristics. Namely, we employed the periodic boundary conditions (PBC) and the fragment approach. We found that the nature of intermolecular interactions plays a significant role in the applicability and performance of both approaches. Specifically, highly polarizable systems such as methanol and ammonia are described by the PBC approach surprisingly well on the MP2 level. For dispersion bound systems such as methane and, to a lower extent, carbon dioxide, we find the fragment approach to be more applicable.

Both approaches have their strengths and weaknesses. The fragment approach generally performs well on systems with dominating pairwise interactions, *i.e.*, systems with negligible 3-body, etc., contributions because of a modest computational cost of pairwise interactions compared to larger clusters. However, this approach suffers from the lack of generally applicable techniques for converging fragment expansion within desired accuracy. In other words, whereas for one system one might be well off considering dimers and trimers within 10 Å, for another system this could be a less plausible approximation. Moreover, the number of trimers that one needs to consider increases significantly with distance which in turn raises not only computer cost but also human time because one must avoid double counting or omitting fragments. Thus, a quite substantial amount of human time is needed to obtain accurate results using the fragment scheme.

On the other hand, the PBC approach is rather simple to set up and a general strategy for converging the results exists, as we showed in our thesis. In principle, lattice energies strongly depend on the k-point set used for the solid and on the cell size used for the reference molecule. For systems consisting of small molecules, such as our four systems, we showed that the lattice energies can be converged to a fairly high precision (less than 1 kJ/mol), provided that enough computer resources are available. In fact, for precise results at the MP2 level, much computer resources are needed to converge the energy of the reference molecule, especially in terms of CPU time and RAM requirements, due to a steep (N^5) scaling of periodic MP2. For example, the MP2 calculations for CO_2 required cca $2 \cdot 10^5$ CPU hours overall and a maximum of 1.5 TB of RAM. In comparison the fragment approach required two orders of magnitude less CPU hours and RAM. Fortunately, however, some progress has been made recently and a quartic scaling periodic MP2 algorithm has been developed [51]. Nevertheless, we showed that for our systems, the RPA algorithm with a better N^3 scaling can be used with advantage in this regard.

As far as we know, a direct comparison of both approaches at the MP2 level has not yet been made. To see whether our set-ups can produce reasonable results, we first performed test caculations on dimers before actual calculations on solids. This was particularly relevant for the PBC approach, where we used pseudopotentials. We sought to assess both precision and accuracy of our results. We checked our cluster calculations on methane and methanol dimers against two benchmark sets and found a very good agreement at both the MP2 and CCSD(T) levels using the explicitly correlated F12 method. Comparing both approaches,

we found that the PBC approach underestimates MP2 correction energies by 2 – 3 % which is a satisfying result given that the convergence of MP2 energies with respect to the cell size was fairly slow, as we discussed in Chapter 2. When performing calculations on solids, the periodic HF and MP2 calculations then allowed us to check if the fragment expansion was sufficiently converged.

One of the biggest problems with the fragment expansion is to know when to stop because there are many small contributions that need to be taken into account. For example, for ammonia cca 98% of the dimer contribution comes from within 9 Å and only 2% from the range of 9 – 18 Å from the central molecule. However, it is usually more difficult to converge the contribution from trimers because of their rather complex behavior. As an illustration, triangle-shaped trimers with symmetric intermolecular bonds usually have a negative interaction energy, thereby contributing to the stronger binding of the solid, while triangle-shaped trimers with asymmetric intermolecular bonds and linear trimers act to the contrary. In fact, there are often quite as many trimers with a negative (binding) contribution as there are trimers with positive (anti-binding) contribution. Consider, for example, the case of CO_2; summing up absolute values of MP2 corrected trimer interaction energies within 10 Å, we found the value of 7 kJ/mol, whereas in reality, their total contribution is only 0.3 kJ/mol. For the Hartree-Fock energies of these trimers, we similarly have 8.8 kJ/mol with absolute values, and –0.7 without them. Therefore, we see that converging these results might be a crucial step towards obtaining precise lattice energies.

It is well known that the mean-field Hartree-Fock method cannot describe molecular solids accurately. Whereas for methanol and ammonia the Hartree-Fock energy constitutes roughly a third of the total CCSD(T) lattice energy, for methane we observed a non-binding (positive) contribution from HF and for carbon dioxide HF gives only cca 12% of the CCSD(T) lattice energy. Thus, we found that the Hartree-Fock method performs more poorly with dispersion-dominated solids than with systems with hydrogen bonds.

Our results from both approaches agree to within 2% at the MP2 level, which is a satisfying result. We considered the basis set incompleteness error, and we usually added a rough estimate of the error emanating from the finite cutoff distance. For all our systems, the uncertainties of MP2 results are below 1 kJ/mol. Thus, at the MP2 level, we produced rather precise lattice energies of our molecular solids. Unfortunately, at the CCSD(T) level, we did not manage to obtain well-converged data for trimers and tetramers of methanol, ammonia, and carbon dioxide, due to a high computational cost. Therefore, we estimate an error of 10% for the Δ[CC–MP2] corrections based on the convergence of HF and MP2 results. We also note that reference data are not of sufficient precision to make a comparison reliable for CCSD(T).

To sum up, we found that both approaches can yield essentially identical results. The PBC approach has advantages for systems with low symmetry although the compute cost is high. Moreover, it takes relatively little time to set up the calculations and obtain converged results. On the other hand, the fragment approach, especially at the CCSD(T) level, can be very time-consuming and it can be very difficult to obtain converged results. Our fragment-based model relying on the use of explicitly correlated F12 methods for calculations of lattice energies could be improved by using force field methods for distant clusters.

Bibliography

[1] J. Bauer, S. Spanton, R. Henry, and J. Quick. Ritonavir: an extraordinary example of conformational polymorphism. *Science*, 283(5398):44–46, 1999.

[2] K. R. Ramya and A. Venkatnathan. Stability and reactivity of methane clathrate hydrates: Insights from density functional theory. *Journal of Physical Chemistry A*, 116:77420–7745, 2012.

[3] C. Cavazzoni, G. L. Chiarotti, and M. Parrinello. Superionic and metallic states of water and ammonia at giant planet conditions. *Science*, 283(5398): 44–46, 1999.

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

[5] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler. Hydrogen bonds and van der Waals forces in ice at ambient and high pressures. *Physical Review Letters*, 107:185701, 2011.

[6] 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. *Journal of Chemical Physics*, 139:154702, 2013.

[7] M. Del Ben, J. Hutter, and J. VandeVondele. Second-order Moller-Plesset perturbation theory in the condensed phase: An efficient and massively parallel gaussian and plane waves approach. *Journal of Chemical Theory and Computation*, 8:4177, 2012.

[8] D. Usvyat. Linear-scaling explicitly correlated treatment of solids: Periodic local MP2-F12 method. *Journal of Chemical Physics*, 139:194101, 2013.

[9] Y. Li, D. Lu, H. Nguyen, and G. Galli. van der Waals Interactions in Molecular Assemblies from First-Principles Calculations. *Journal of Physical Chemistry A*, 114:1944, 2010.

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

[11] 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*, 106:153003, 2011.

[12] 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. *Journal of Chemical Physics*, 132(154104), 2010.

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

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

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

[16] A. M. Reilly and A. Tkatchenko. Understanding the role of vibrations, exact exchange, and many-body van der waals interactions in the cohesive properties of molecular crystals. *Journal of Chemical Physics*, 139(024705), 2013.

[17] S. L. Price. Computer crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. *Accounts of Chemical Research*, 42:117, 2009.

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

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

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

[21] O. Bludský, M. Rubeš, and P. Soldán. Ab initio investigation of intermolecular interactions in solid benzene. *Physical Review B*, 77:092103, 2008.

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

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

[24] C. Červinka, M. Fulem, and K. Ružička. CCSD(T)/CBS fragment-based calculations of lattice energy of molecular crystals. *Journal of Chemical Theory and Computation*, 144, 2016.

[25] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Revised. Dover Publications, New York, 1996. ISBN 0-486-69186-1.

[26] F. Jensen. *Introduction to Computational Chemistry*. Second ed. Wiley, 2007. ISBN 0470058048.

[27] F. Neese, F. Wennmohs, and A. Hansen. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *Journal of Chemical Physics*, 130(114108), 2009.

[28] G. Schmitz, C. Hättig, and D. P. Tew. Explicitly correlated PNO-MP2 and PNO-CCSD and their application to the s66 set and large molecular systems. *Physical Chemistry Chemical Physics*, 16:22167–22178, 2014.

[29] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schutz. Molpro: a general-purpose quantum chemistry program package. *WIREs Computational Molecular Science*, 2:242–253, 2012.

[30] W. Kutzelnigg and W. Klopper. Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. i. general theory. *Journal of Chemical Physics*, 94:1985, 1991.

[31] H.-J. Werner, T. B. Adler, and F. R. Manby. General orbital invariant MP2-F12 theory. *Journal of Chemical Physics*, 126(164102), 2007.

[32] L. Kong, F. A. Bischoff, and E. F. Valeev. Explicitly correlated R12/F12 methods for electronic structure. *Chemical Reviews*, 112:75–107, 2012.

[33] T. Helgaker, A. Halkier, P. Jørgensen, W. Klopper, and J. Olsen. Basis-set convergence of the energy in molecular Hartree–Fock calculations. *Chemical Physics Letters*, 302:437–446, 1999.

[34] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Physical review B*, 47, 1993.

[35] G. Kresse and J. Hafner. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Physical review B*, 49, 1994.

[36] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54, 1996.

[37] M. Marsman, A. Grüneis, J. Paier, and G. Kresse. Second-order Møller–Plesset perturbation theory applied to extended systems. I. structural and energetic properties. *Journal of Chemical Physics*, 130(184103), 2009.

[38] P. E. Blöchl. Projector augmented-wave method. *Physical review B*, 50: 17953, 1994.

[39] G. Kresse and J. Joubert. From ultrasoft pseudopotentials to the projector augmented wave method. *Physical review B*, 59, 1999.

[40] F. Gygi and A. Baldereschi. Self-consistent Hartree-Fock and screened-exchange calculations in solids. *Physical review B*, 34:309–319, 1986.

[41] 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: 205119, 2006.

[42] A. Grüneis, M. Marsman, and G. Kresse. Second-order Møller–Plesset perturbation theory applied to extended systems. II. within the projector-augmented-wave formalism using a plane wave basis set. *Journal of Chemical Physics*, 133(074107), 2010.

[43] M. Kaltak, J. Klimeš, and G. Kresse. Cubic scaling algorithm for the random phase approximation: Self-interstitials and vacancies in si. *Physical review B*, 90, 2014.

[44] P. Jurečka, J. Šponer, J. Černý, and P. Hobza. Benchmark database of accurate (MP2 and CCSD(T) CBS limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Physical Chemistry Chemical Physics*, 8, 2006.

[45] J. Řezáč, K. E. Riley, and P. Hobza. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *Journal of Chemical Theory and Computation*, 7(2427), 2011.

[46] S. Grazulis, D. Chateigner, R. T. Downs, A. F. T. Yokochi, M. Quiros, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck, and A. Le Bail. Crystallography open database - an open-access collection of crystal structures. *Journal of Applied Crystallography*, 42:726, 2009.

[47] B. H. Torrie, S. X. Weng, and B. M. Powell. Structure of the α-phase of solid methanol. *Molecular Physics*, 67:575–581, 1989.

[48] A. Simon and K. Peters. Single-crystal refinement of the structure of carbon dioxide. *Acta Crystallographica B*, 36:2750, 1980.

[49] R. Boese, N. Niederprüm, D. Bläser, A. Maulitz, M. Y. Antipin, and P. R. Mallinson. Single-crystal structure and electron density distribution of ammonia at 160 K on the basis of X-ray diffraction data. *Journal of Physical Chemistry B*, 101:5794, 1997.

[50] C. Červinka and M. Fulem. State-of-the-art calculations of sublimation enthalpies for selected molecular crystals and their computational uncertainty. *Journal of Chemical Theory and Computation*, In Press, 2017.

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

List of Figures

List of Tables

List of Abbreviations

PBC	periodic boundary conditions
HF	Hartree-Fock
MP2	Second order Møller-Plesset perturbation theory
CC	coupled cluster
CI	configuration interaction
DFT	density functional theory
RPA	random phase approximation
PAW	projector-augmented-wave
AVXZ	aug-cc-pVXZ
CBS	complete basis set
BSSE	basis set superposition error
BSIE	basis set incompleteness error
Me	methane
MeOH	methanol

Attachments