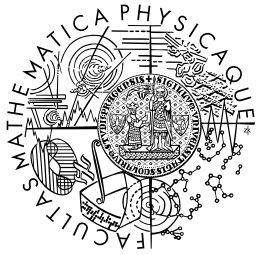


FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Adam Mendl

Calculation of Intermolecular Interactions

Department of Chemical Physics and Optics

Supervisor of the bachelor thesis: Mgr. Jiří Klimeš, Ph.D.
Study programme: Physics
Study branch: Physics (B0533A110001)

Prague 2024

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

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

In ..... date ....

Author's signature

I would like to thank Mgr. Jiří Klimeš, PhD. for introducing me to accurate energy methods and to multipole methods. Also, I am grateful for pleasant consultations where he answered all of my questions regarding many body physics. I have learned a lot. Finally, I cannot forget my family – Ludmila, Vladimír and David – who supported me the whole time.

Název práce: Výpočty mezimolekulárních interakcí

Author: Adam Mendl

Katedra: Katedra chemické fyziky a optiky

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

Abstract: Predikce vlastností a struktury molekulárních krystalů (krystalů skládajících se z molekul vázaných nekovalentními vazbami) z ab-inito výpočtů je náročný problém. Vazebné energie takovýchto struktur vyžadují použití přesných ab-initio metod, které jsou výpočetně extrémně náročné. Potenciály, které lze nafitovat na ab-initio data nebo jejichž parametry lze získat z předem vypočtené matice hustoty, lze použít k popisu interakcí vzdálených molekul, čímž lze snížit náročnost celého výpočtu. Jednou z metod zaměřených na popis elektrostatických příspěvků je distribuovaný multipólový rozvoj rozvádějící matici hustoty do sférických tenzorů kolem několika bodů na molekule. Vyvinuli jsme program, který dokáže získat koeficienty multipólového rozvoje ze dvou různých kódů integrujících matici hustoty, vypočítat elektrostatické příspěvky mezimolekulární energie a otáčet multipólový rozvoj. Toto lze provést do libovolného řádu. Dále studujeme vlastnosti vypočítaných koeficientů multipólového rozvoje v závislosti na ab-initio metodě, bázi matice hustoty a kódu, který byl použit pro integraci matice hustoty. Otestovali jsme predikci elektrostatické interakce na molekolárních krystalech s různými fyzikálně-chemickými vlastnostmi – methanol, metan, amoniak a oxid uhličitý. Dále jsme se také pokusili nafitovat koeficienty jednoho multipólového rozvoje na molekule na předpočítaných datech interakčních energiích dimerů. Abychom otestovali možnost získávání distribuovaných vlastností z existujících datasetů, vyzkoušeli jsme proložení jednoduchého modelu korelační energie pomocí klasických fitovacích postupů i technik ze strojového učení.

Klíčová slova: mezimolekulární interakce, ab-initio, kvantová mechanika, multipólová analýza

Title: Calculation of Intermolecular Interactions

Author: Adam Mendl

Department: Department of Chemical Physics and Optics

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

Abstract: Prediction of properties and structure of molecular solids (crystals comprising molecules bonded by non-covalent interactions) from ab-inito calculation is a challenging problem. Binding energies of such structures require usage of precise ab-initio methods, which are computationally extremely demanding. Potentials which can be fitted on ab-initio data, or whose parameters can be obtained from a precalculated density matrix, can be used to reduce the number of computations by describing interactions of distant molecules. One method aiming to describe electrostatic contributions is distributed multipole expansion, which expands a density matrix into spherical tensors around several points on a molecule. We have developed a program that can obtain coefficients of multipole expansion from two different codes integrating the density matrix, calculate electrostatic contributions to intermolecular energy, and rotate multipole expansion. This can be done up to an arbitrary order of expansion. We study properties of calculated coefficients depending on ab-initio method, basis of the density matrix, and the code that was used for the integration of the density matrix. We test the prediction of electrostatic energy on dimer datasets of molecular solids with different physical/chemical properties – methanol, methane, ammonia, and carbon dioxide. Furthermore, we also tried to fit the coefficients of single multipole expansion on dimer dataset. To test the possibility of description of distributed properties, we tried fitting a simple model of correlation energy by using classical fitting procedures as well as techniques from machine learning.

Keywords: intermolecular interactions, ab-initio, quantum mechanics, multipole analysis

Contents

1. Introduction

1.1 Molecular solids

Molecular solids are materials consisting of molecules bonded by forces (van der Waals, dipole–dipole, London dispersion[1]) which are typically much weaker than metallic or covalent bonds occurring in solids. Molecular solids include, for example, hydrates, solvates, salts, and materials containing molecules trapped inside cages formed by other molecules (so-called clathrate hydrates).

Materials from weakly bonded molecules are of giant importance in many branches of science, technology, and especially in pharmaceutical engineering. To mention at least few applications, porous molecular solids can be used as specialized sieves [2] and porous molecular liquids [3] with similar applications as porous molecular solids are gaining increasing traction among researches. Some atmospheric layers of several giant planets consist mostly of methane, ammonia, or water, and those chemicals influence their chemical and (astro)physical properties [4]. Natural gas hydrates present in marine and permafrost environments have been extensively studied as an alternative source of energy by research groups supported by investments of several governments, as well as private entities [5].

Since the structure of molecular solids is partially determined by weak bonds, they often exhibit conformational polymorphic behavior, which means that they can exist with various crystal-packing motifs. The energy differences between polymorphs are usually a few kJ/mol or less [6]. This leads to a pallet of quite interesting phenomena. For example, freezing water is still not well understood (since ice has at least 19 different phases known today) and the intermolecular forces within ice are still an active field of research [7, 8, 9, 10]. Conversely, polymorphism leads to many issues regarding practical applications. Probably the most obstacles are present in the field of pharmaceutical engineering, where the bioavailability of drugs has to be diligently checked. A famous example of compromised bioavailability is the HIV drug ritonavir, which was introduced to the market in 1998. Supplies of this drug were threatened by a new, in that time unknown, form of the crystal. This ultimately led to removal from the market and huge financial losses [11].

In general, obtaining precise binding energies by direct measurement is non-trivial. Errors in the experimentally obtained values can be of the same magnitude as values obtained by quantum mechanical calculations. Measured values also need some corrections (for example for thermal and zero-point effects) which increase relative uncertainty of the final results [12]. For this thesis, relevant measurements obtained from experimental sublimation enthalpies can be found in [13] (ammonia, carbon dioxide) and in [14] (methanol, ammonia, carbon dioxide). It may be intriguing to mention that, to our knowledge, experimentally obtained

[1] London dispersion arises from temporary fluctuations in electron distribution within molecules due to the constant movement of electrons. The movement of electrons in the two interacting molecules becomes correlated. Therefore, dipole-dipole like interaction is introduced (it favors lower energy configurations). Since the correlation between the two electrons becomes stronger when the molecules are closer to each other, the resulting force has an attractive character [1].

values of binding energies of methane were not yet published. Moreover, some rather extreme conditions with high pressure and temperature usually cannot be replicated in the laboratory at all. To make the situation even more complicated, molecular solids might find applications in the quickly growing field of organic (opto)electronics, where not only the ground state, but also excited state studies have to be performed [15].

1.2 Computational techniques for molecular solids

The problems with experimental investigation of molecular solids make theoretical approaches to binding energies essential. The field of precise energy methods for big weakly bonded systems has been expanded significantly by considerable advances in computational power, as well as by the development of more accurate and more optimized ab-initio quantum chemistry methods. Calculations of binding energies can be done in two ways: so-called fragmented approach utilizing many body expansion or by employing periodic boundary conditions [6].

In the fragmented approach, one central molecule is chosen, and the whole energy is summed in terms of 2-body, 3-body, 4-body etc. contributions [16]. This approach would, to obtain exact energies, require summing infinitely many terms, which is of course impossible. However, it suffices to sum a finite number of terms, which has to be selected and diligently analyzed according to the required precision of converged value [17]. Since summing and analyzing 2-body terms is usually straightforward, the fragmented approach works best on systems where pairwise interaction is prevalent. The number of trimers that one needs to evaluate grows significantly with the system size, and the analysis of convergence of these terms can be nontrivial. The steep growth of the number of trimers, tetramers, etc. also means that effects such as polarizability, which quickly vanish regarding the distance, start to become significant. This, of course, puts pressure on the accuracy of used computational methods as well as on the computational resources that have to be dedicated for obtaining reasonably precise energies.

Many ab-initio methods were tested for molecular solids in the fragmented approach [6]. Benchmark calculations are typically produced by coupled clusters, which is the golden standard of accurate chemical methods for mid-size systems. However, since the coupled clusters have massive computational cost, it is typically not feasible to evaluate the energy of the whole crystal using this particular method and different algorithms have to be employed (often after checking, that they give similar results as coupled clusters benchmarks). Density Functional Theory (DFT) can, with careful treatment of exchange energy and nontrivial dispersion corrections (see [18] for example), lead to accurate results for dimer contributions. However, widely used approximations have large errors for three-body and higher-body contributions [19, 20]. A little more computationally expensive method, called Random Phase Approximation, is known to underestimates binding energies by even 20%. On the other hand, it is usually more accurate than DFT and singles corrections lead to significant improvement for some systems [21]. Møller-Plesset perturbation theory can in some cases be used for accurate results when summation of big number of terms is needed [17]. Møller-Plesset perturba-

tion theory has problems with the description of highly polarizable molecules.

Using methods within periodic boundary conditions might seem as a better alternative to many-body expansion; yet, the most accurate algorithms (coupled clusters) cannot be used within these settings due to significant computational cost (except for the simplest of systems [22]). DFT is regularly used within periodic boundary conditions (however, its results have to be checked). Successful periodic implementations of second order Møller-Plesset perturbation theory (see for example [23]) are rather recent with fairly unfavorable scaling regarding the system size $\mathcal{O}(N^5)$, and the first implementation with better yet not ideal quartic scaling was introduced no later than in 2017 [24]. Random Phase Approximation can also be used in the periodic settings; however, usually, not by applying a self-consistent procedure directly but by building upon the calculations of DFT [25].

1.3 Our work

In this work, we will study the possibility to apply distributed multipole expansion for description of electrostatic effects of dimers. Even though Distributed multipole analysis (DMA) is quite an old concept [26], the study, of how well the electrostatic interaction of molecular solids can be predicted was not done in the systematic way for rather recent calculations. Furthermore, the aim of this work is to develop required software to simplify usage of this technique. We are also interested in the investigation of a few issues that frequently arise when employing distributed multipole analysis in practice. What we mean by that is, for example, poor documentation of software packages and availability of methods in quantum chemistry ab-initio codes. Furthermore, during this work, we will find that reliability and stability of coefficients of multipole expansion obtained from programs integrating density matrices has a crucial significance for precision. We will also study, how to deal with orientations and rotations of structures and their multipole expansions. The DMA is usually done for big molecules, where the electrostatic field may be well described only by charges and multipoles of low order, which is a wholly different situation from systems that we wish to work with in this text. For smaller molecules, accurately calculated higher order multipoles will be required. This will turn out to be rather challenging. We will also test, how well the parameters of the multipole expansion can be fitted from calculated dimer data. The ultimate goal is to develop software, which could accurately calculate distant electrostatic dimer contributions. This would lead to significant speed up for evaluation of binding energies of molecular crystals. When the binding energies obtained by ab-initio methods are small, the numerical errors become significant. Therefore, DMA can probably give more precise distant electrostatic contributions than ab-initio calculations in some cases.

In the second chapter, we will introduce some ab-initio methods into greater details and briefly mention several others. Subsequently, we will devote the third chapter to the description of distributed multipole expansion. Finally, the last chapter deals with results regarding stability and convergence of expansion as well as comparison with rather recent results obtained by ab-initio quantum mechanical calculations of methane, methanol, ammonia, and carbon dioxide. Possibilities how to fit multipole expansion are explored. The atomic units are applied.

2. Ab initio methods

In electronic structure theory, we aim to solve the stationary Schrödinger equation

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\ , \tag{2.1}$$

where the Hamiltonian takes the form

$$\hat{H} = \underbrace{\sum_i \frac{1}{2m}\hat{\mathbf{p}}_i^2}_{\text{kinetic energy}} + \underbrace{\frac{1}{2}\sum_{i\neq i'}\frac{1}{|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_{i'}|}}_{\text{electron}\,-\,\text{electron repulsion}} - \underbrace{\sum_{ij}\frac{Z_j}{|\hat{\mathbf{x}}_i - \mathbf{R}_j|}}_{\text{electron}\,-\,\text{nuclear attraction}} \tag{2.2}$$

and where $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{p}}_i$ are operators of position and momentum of electrons and $\mathbf{R}_j$ and Z_j are positions and charges of nuclei.

During calculations throughout this work, we assume that nuclei are fixed and positions R_j are known. Moreover, we ignore any nuclear quantum effects (quantum effects of hydrogen can contribute to energies of some systems; however, in our case other sources of error have much greater significance). Therefore, the structure of (2.2) is simple. Even in this rather straightforward situation, the solution of equation (2.1) with Hamiltonian (2.2) is not treatable analytically. This is where so-called ab-initio quantum chemistry methods [27] (meaning from first principles) are used to approximately obtain eigenstates and eigenvalues of the Hamiltonian (2.2). After the groundstate of the electronic problem and its energy E are found, the total ground state energy of the system can be written as

$$E_{\text{total}} = E + \underbrace{\frac{1}{2}\sum_{j\neq j'}\frac{Z_j Z_{j'}}{|\mathbf{R}_j - \mathbf{R}_{j'}|}}_{\text{nuclear}\,-\,\text{nuclear repulsion}}\ . \tag{2.3}$$

In the following parts, we will discuss several methods how to deal with (2.2). We will start with the variational method and use it to obtain a Self Consistent Field (SCF) procedure based on Hartree-Fock ansatz. Subsequently, we will introduce post-HF methods such as Configuration Interaction (CI), Møller-Pleset perturbation theory (MP) and Coupled Clusters (CC). We will finish this chapter with Random Phase Approximation (RPA) and by discussing basis sets, computational scaling and explicitly correlated methods. But before that, let us briefly introduce other algorithms for quantum chemistry.

1. Quantum Monte Carlo is a method based on sampling of the wave function. The most utilized variant is so-called diffusion Monte Carlo, which uses Green functions. The striking property of this method is that for bosons the algorithm scales polynomially in the system size, but for fermions the scaling is actually exponential – this is resolved by so-called Fixed node Monte Carlo [28].

2. Density Functional Theory with many functionals is probably the most influential computational technique in many-body physics. Today, there exists over 200 density functionals with applications to different chemical systems [29, 30]. DFT methods are also popular for describing time evolution [31] or systems in non-equilibrium states [32]. We also cannot forget its

applicability for nuclei [33] and possibly in the future for neutron stars [34]. Density Functional Theory describes many systems better than Restricted Hartree-Fock (RHF), since DFT approximately includes electron correlations.

3. Simulations of quantum systems were supposed to be the first applications of quantum computers. Despite massive progress in both theoretical aspects and physical implementations [35], meaningful calculations with higher performance than conventional computers were obtained only for specific quantum systems [36] (and some of these claims were later heavily criticized [37]). Current experiments with simulating electronic structure on quantum computers did not produce significantly better results than conventional ones and needed prepossessing. Furthermore, it is not possible to simulate particularly large systems because of substantial noise occurring, since a higher than expected number of gates is required [38]. Nevertheless, small systems were already simulated on conventional computers emulating quantum circuits. For methods, see [39] or [40].

2.1 Variational method

Since variational principle is backbone of quantum chemistry calculations and is presented in every textbook on this topic, we will try to introduce this technique in precise mathematical terms. Let us have hilbert space $\mathcal{H}$ with hermitian hamiltonian $\hat{H}$ with spectrum bounded by highest lower bound E_0 (this is obvious condition for any physical system) and discrete spectrum[1]. Furthermore, let $\{\lambda_i\}_{i>1}$ be eigenvalues corresponding to orthogonal eigenvectors $\{\psi_{\lambda_i}\}_{i>1}$ that are basis vectors of eigenspace of $\hat{H}$.

We can summarize these relations by

$$\begin{aligned} \langle\psi_{\lambda_i}|\psi_{\lambda_j}\rangle &= \delta_{ij}\,, \\ \hat{H}\,|\psi_{\lambda_i}\rangle &= \lambda_i\,|\psi_{\lambda_i}\rangle\ , \\ \min_{i>1}\lambda_i &= E_0\,. \end{aligned} \tag{2.4}$$

Using these relations, we can obtain for any $|\psi\rangle \in \mathcal{H}$

$$\langle\psi|\,\hat{H}\,|\psi\rangle = \sum_{i,j}\langle\psi|\psi_{\lambda_i}\rangle\,\langle\psi_{\lambda_i}|\,\hat{H}\,|\psi_{\lambda_j}\rangle\,\langle\psi_{\lambda_j}|\psi\rangle = \sum_i \lambda_i|\,\langle\psi_{\lambda_i}|\psi\rangle\,|^2 \geq E_0\,\langle\psi|\psi\rangle\ , \tag{2.5}$$

which can be rewritten into the well-known equation

$$\frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} \geq E_0\,. \tag{2.6}$$

[1] Another reasonable expectation of bound states in quantum chenistry. If we wanted to incorporate continuous spectrum into consideration, we would have had to simply rewrite equation (2.4) into the form $\langle\psi_i|\psi_j\rangle = \delta(i-j)$ where the Kronecker delta is replaced by delta function. On the other hand, this step would make all calculations discussed here impossible and the power of variational technique would be basically gone.

Now, we will take a subspace $\mathcal{A} \subset \mathcal{H}$ of dimension N and parameterize it with the set $\{\alpha_i\}_{i=1}^{N} \subset \mathbb{C}^N$. By varying

$$\varepsilon(\alpha_1, \ldots, \alpha_N) = \langle \psi(\alpha_1, \ldots, \alpha_N) | \hat{H} | \psi(\alpha_1, \ldots, \alpha_N) \rangle \tag{2.7}$$

over all wave functions $\psi(\alpha_1, \ldots, \alpha_N) \in \mathcal{A}$ with constraint

$$\langle \psi(\alpha_1, \ldots, \alpha_N) | \psi(\alpha_1, \ldots, \alpha_N) \rangle = 1 \tag{2.8}$$

we should obtain the ground state $|\psi_0\rangle \in \mathcal{A} \subset \mathcal{H}$. We should vary over $\mathcal{H}$, but since varying over infinitely dimensional vector space is hard, we have reduced our problem to a smaller ansatz $\mathcal{A}$. To be able to solve this variational problem, the ansatz $\mathcal{A}$ has to overlap with the subspace of all ground states $|\psi_0\rangle$. If this condition doesn't hold, we obtain only upper limit to ground state energy, and we have to choose a different ansatz $\mathcal{A}$.

One possibility how to make variational method work better is to iteratively expand the ansatz $\mathcal{A}$ (for example, see [41]). However, even in this situation it has been shown on toy model of so-called spiked harmonic oscillator[2], that the process $\varepsilon(\alpha_1, \ldots, \alpha_N) \to E_0$ does not have to converge to the actual ground state wave function $|\psi_0\rangle$. The problem lies in the mathematical properties of infinitely dimensional vector spaces where the limit $\lim_{n\to\infty} \psi_n$ does not have to make sense and therefore energy given by (2.7) converges only slowly or not at all (for example of really slow convergence of ground state energy and nonsensical relations for ground state see conclusion section in [41]).

If we obtain the ground state $|\psi_0\rangle$, it is possible to find excited states by variational principle as well. We choose ansatz $\mathcal{B} \subset \mathcal{H}$ of dimension N' that is orthogonal to the space of all ground states $|\psi_0\rangle$. Similarly, as for the ground state calculations, we parametrize $\mathcal{B}$ with set $\{\beta\}_{i=1}^{N'}$ and vary over these parameters. If $\mathcal{B}$ has nontrivial intersection with the first excited state, our approach might converge to the wave function corresponding to the first excited state. In theory, it should be possible to find all energies and all eigenstates of any system using this method. However, the higher the order of excited state, the less precise the practical calculations are. Moreover, various variational methods have been proposed for the direct computation of excited states (even by the mean-field approach) [42, 43, 44].

Variational principle stands behind Variational Monte Carlo, which employs complicated trial wave functions to obtain ground state. The ansatz $\mathcal{A}$ may be parametrized by a huge number of parameters, or it can be encoded into a neural network – this is known as neural network quantum states [45]. Another variational technique is Density Matrix Renormalization Group [46] and Time Evolving Block Decimation, which is the most efficient method for many one dimensional systems (especially systems with low number of entangled particles).

2.2 Restricted Hartree-Fock closed shell calculations and Roothan equations

The simplest approach to ab-initio electronic structure theory is to create a mean-field operator describing electrostatic interaction between electrons. This then

[2]Hamiltonian of this system is given by $H = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x^2 + \frac{l(l+1)}{x^2} + \lambda|x|^{-\alpha}$ [41].

allows an efficient iterative procedure, when the mean-field operator is evaluated by an "older" guess for the solution, and it is used to obtain "newer" one. In this section, we mostly use Chapter 3 from [47] as a point of reference for general equations and the derivation of the restricted method is taken from [48].

The main postulate of the system of fermions is the condition of antisymmetric wave function with respect to the interchange of two particles. This leads us to write the wave function as determinant (so-called Slater determinant) of for now orthonormal spin orbitals $\psi_i(\boldsymbol{x}_j)$

$$|\Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} |\psi_1(\boldsymbol{x}_1)\rangle & |\psi_2(\boldsymbol{x}_1)\rangle & \ldots & |\psi_N(\boldsymbol{x}_1)\rangle \\ |\psi_1(\boldsymbol{x}_2)\rangle & |\psi_2(\boldsymbol{x}_2)\rangle & \ldots & |\psi_N(\boldsymbol{x}_2)\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |\psi_1(\boldsymbol{x}_N)\rangle & |\psi_2(\boldsymbol{x}_N)\rangle & \ldots & |\psi_N(\boldsymbol{x}_N)\rangle \end{vmatrix}. \tag{2.9}$$

The antisymmetry properties as well as the Pauli exclusion principle then hold directly from the properties of determinant. We may define function χ as a main diagonal of the Slater determinant

$$|\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\rangle = |\psi_1(\boldsymbol{x}_1)\ldots\psi_N(\boldsymbol{x}_N)\rangle \tag{2.10}$$

and the anti symmetrization operator, which will be used to obtain $|\Psi\rangle$ from $|\chi\rangle$

$$\hat{A} = \frac{1}{\sqrt{N!}}\sum_{p=1}^{N-1}(-1)^p\hat{P}_p\,, \tag{2.11}$$

where $\hat{P}_p$ is the permutation operator. The antisymetrization operator (antisymmetrizer) has obvious properties[3]

$$\hat{A} = \hat{A}^\dagger\,, \quad \hat{A}\hat{A} = \sqrt{N!}\hat{A}\,, \quad \left[\hat{A}\,,\hat{H}\right] = 0\,, \tag{2.12}$$

where $\hat{H}$ is any observable (hermitian operator). Now, we may write

$$\begin{aligned} |\Psi(\boldsymbol{x}_1\ldots,\boldsymbol{x}_n)\rangle &= \frac{1}{\sqrt{N!}}\sum_{i=0}^{N-1}(-1)^i\hat{P}\,|\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\rangle \\ &= \frac{1}{\sqrt{N!}}\left(\hat{I} + \sum_{n=1}^{N}\sum_{i_1\ldots i_n}(-1)^n\hat{P}_{i_1\ldots i_n}\right)|\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\rangle \\ &= \hat{A}\,|\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\rangle\ . \end{aligned} \tag{2.13}$$

We now introduce operators $\hat{h}_i$ and $\hat{g}_{ij}$, where $\hat{h}_i$ is the operator of kinetic energy and potential energy of electron inside the field of nuclei and $\hat{g}_{ij}$ is the repulsion of electrons. Using this notation, the Hamiltonian (2.2) may be separated into one- and two-electron operators as

$$\hat{H} = \sum_{i=1}^{N}\hat{h}_i + \sum_{i=1}^{N}\sum_{j>i}^{N}\hat{g}_{ij}\,. \tag{2.14}$$

[3]For the first property, it is enough to realize, that if π is a permutation, then π and π^{-1} have the same parity. The property then holds if the permutation is a unitary transformation, which obviously is true. The second equation holds from the definition of projector operator. The last relation is obvious, since if the antisymtrizer did not commute with a hermitian operator corresponding to the observable $\hat{H}$, the indistinguishability of particles would not hold.

We can derive matrix elements of the single electron operator $\hat{h}_i$ using the orthonormality of spin orbitals [47]

$$\begin{aligned}\langle\chi|\hat{h}_i|\chi\rangle &= \langle\psi_1(\boldsymbol{x}_1)\dots\psi_N(\boldsymbol{x}_N)|\hat{h}_i|\psi_1(\boldsymbol{x}_1)\dots\psi_N(\boldsymbol{x}_N)\rangle \\ &= \langle\psi_1(\boldsymbol{x}_1)|\psi_1(\boldsymbol{x}_1)\rangle\dots\langle\psi_i(\boldsymbol{x}_i)|\hat{h}_i|\psi_i(\boldsymbol{x}_i)\rangle\dots \\ &\quad \langle\psi_N(\boldsymbol{x}_N)|\psi_N(\boldsymbol{x}_N)\rangle \\ &= \langle\psi_i(\boldsymbol{x}_i)|\hat{h}_i|\psi_i(\boldsymbol{x}_i)\rangle\ .\end{aligned} \tag{2.15}$$

We may also see that all matrix elements involving permutations are identically zero. Regarding the two-electron operator, only identity and two electron permutations can lead to nonzero matrix elements. Higher electron permutations are identically zero (this is a special example of Slater-Condon rules [49]). Direct calculations of two-electron operator matrix elements give us [47]

$$\begin{aligned}\langle\chi|\hat{g}_{ij}|\chi\rangle &= \langle\psi_1(\boldsymbol{x}_1)\dots\psi_N(\boldsymbol{x}_N)|\hat{g}_{ij}|\psi_1(\boldsymbol{x}_1)\dots\psi_N(\boldsymbol{x}_N)\rangle \\ &= \langle\psi_1(\boldsymbol{x}_1)|\psi_1(\boldsymbol{x}_1)\rangle\dots\langle\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)|\hat{g}_i j|\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)\rangle\dots \\ &\quad \langle\psi_N(\boldsymbol{x}_N)|\psi_N(\boldsymbol{x}_N)\rangle \\ &= \langle\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)|\hat{g}_i j|\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)\rangle = J_{ij}\,,\end{aligned} \tag{2.16}$$

where J_{ij} represents coulomb interactions of electrons and [47]

$$\begin{aligned}\langle\chi|\hat{g}_{ij}|\hat{P}_{ij}\chi\rangle &= \langle\psi_1(\boldsymbol{x}_1)\dots\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)\dots\psi_N(\boldsymbol{x}_N)|\,\hat{g}_{ij} \\ &\quad |\psi_1(\boldsymbol{x}_1)\dots\psi_i(\boldsymbol{x}_j)\psi_i(\boldsymbol{x}_j)\dots\psi_N(\boldsymbol{x}_N)\rangle \\ &= \langle\psi_i(\boldsymbol{x}_i)\psi_j(\boldsymbol{x}_j)|\,\hat{g}_{ij}\,|\psi_i(\boldsymbol{x}_j)\psi_i(\boldsymbol{x}_j)\rangle = K_{ij}\,.\end{aligned} \tag{2.17}$$

Matrix elements K_{ij} are usually called the exchange integrals.

The variational principle states, that the best wave function minimizes the energy functional. We may write the energy functional of the Slater determinant (which depends on the spin orbitals $\psi_i(\boldsymbol{x})$) as

$$E[\{\psi_i\}_{i=1}^N] = \langle\Psi|\hat{H}|\Psi\rangle = \sqrt{N!}\,\langle\chi|\hat{H}|\hat{A}\chi\rangle = \sum_p (-1)^p\,\langle\chi|\hat{H}|\hat{P}\chi\rangle\ . \tag{2.18}$$

Using results from above, we may rewrite the energy of (2.9) as

$$E[\{\psi_i\}_{i=1}^N] = \sum_i^N h_i + \frac{1}{2}\sum_i^N\sum_j^N (J_{ij} - K_{ij})\,, \tag{2.19}$$

where the minus sign follows from equation (2.11) – see derivation in [47]. We continue by defining the Coulomb, exchange and Fock operator [47]

$$\begin{aligned}\hat{J}_i\,|\psi_j(\mathbf{r}_2)\rangle &= \langle\psi_i(\mathbf{r}_1)|\hat{g}_{12}|\psi_i(\mathbf{r}_1)\rangle\,|\psi_j(\mathbf{r}_2)\rangle\ , \\ \hat{K}_i\,|\psi_j(\mathbf{r}_2)\rangle &= \langle\psi_i(\mathbf{r}_1)|\hat{g}_{12}|\psi_j(\mathbf{r}_1)\rangle\,|\psi_i(\mathbf{r}_2)\rangle\ , \\ \hat{F}_i &= \hat{h}_i + \sum_j^N (\hat{J}_j - \hat{K}_j)\,.\end{aligned} \tag{2.20}$$

We minimize the energy functional with constraint $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. We use the method of Lagrange multipliers ε_{ij}

$$\delta E - \delta\sum_{i=1}^N\sum_{j=1}^N (\varepsilon_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij})) = 0\,. \tag{2.21}$$

Using relations (2.14) and (2.20), we rewrite equation (2.21) into the form [47]

$$\sum_{i=1}^{N} \langle \delta\psi_i | \hat{F}_i | \psi_i \rangle - \sum_{ij}^{N} \varepsilon_{ij} \langle \delta\psi_i | \psi_j \rangle + \sum_{i=1}^{N} \langle \delta\psi_i | \hat{F}_i | \psi_i \rangle^* - \sum_{ij}^{N} \varepsilon_{ij} \langle \delta\psi_j | \psi_i \rangle^* = 0 \,. \quad (2.22)$$

Since the first two and last two terms must cancel in the equation (2.22), we obtain Hartree–Fock equations as [47]

$$\hat{F}_i \psi_i = \sum_{j=1}^{N} \varepsilon_{ij} \psi_j \,. \quad (2.23)$$

From the equation (2.22), it also follows that ε_{ij} are matrix elements of hermitian matrix.

Up to now, the Hartree-Fock method was treated in its full generality. However, we will now restrict ourselves to closed-shell calculations by simplification, which leads to the RHF method announced in the title of this chapter (a brief comment on open shell systems is present at the end of this section). We assume that N electrons are distributed in $\frac{N}{2}$ spatial orbitals (meaning that these orbitals are doubly occupied by two electrons with different spin and the rest is empty). Hence, we change the bounds of the sums [48]

$$E = 2 \sum_{i=1}^{N/2} h_i + \sum_{i=1}^{N/2} \sum_{j \geq i}^{N/2} \left(2\hat{J}_{ij} - \hat{K}_{ij} \right) \,. \quad (2.24)$$

The coefficient 2 in front of the exchange integral K_{ij} is missing, since the exchange interaction is present only between electrons with parallel spins [48]. We expand functions $|\psi_i\rangle$ by set of basis functions $|\phi_\nu\rangle$ (for now generators of the relevant Hilbert space) as

$$\psi_i(\mathbf{r}) = \sum_{\nu} c_{\nu i} \phi_\nu(\mathbf{r}) \,, \quad (2.25)$$

which leads to Hartree-Fock equation in the form

$$\hat{F}(\mathbf{r}) \sum_{\nu} c_{\nu i} \phi_\nu(\mathbf{r}) = \varepsilon_i \sum_{\nu} c_{\nu i} \phi_\nu(\mathbf{r}) \,. \quad (2.26)$$

The next obvious step is to multiply the equation (2.26) by $\phi^*_\mu(\mathbf{r})$ from the left

$$\sum_{\nu} c_{\nu i} \phi^*_\mu(\mathbf{r}) \hat{F}(\mathbf{r}) \phi_\nu(\mathbf{r}) = \varepsilon_i \sum_{\nu} c_{\nu i} \phi^*_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \,. \quad (2.27)$$

By integrating both sides of the equation (2.27) over all space and writing [48]

$$F_{\mu\nu} = \int_{\mathbb{R}^3} \phi^*_\mu(\mathbf{r}) \hat{F}(\mathbf{r}) \phi_\nu(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} \quad (2.28)$$

and [48]

$$S_{\mu\nu} = \int_{\mathbb{R}^3} \phi^*_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \, \mathrm{d}^3\mathbf{r} \,, \quad (2.29)$$

we obtain in the matrix notation the famous Roothan (or Roothaan-Hall) equations [47, 48]

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \,, \quad (2.30)$$

where **F** is called Fock matrix, **C** (with the matrix elements $c_{\nu i}$) is the density matrix and **S** is the overlap matrix. This constitutes a generalized eigenvalue problem, conventionally addressed by orthogonalizing the basis functions.

From specified positions and charges of nuclei, we may calculate the overlap matrix from a specified basis set. If the particular basis set is orthonormal, the overlap matrix becomes a diagonal matrix. When the overlap matrix is known, an initial guess of the density matrix has to be obtained. Finally, we may repeat the two following steps (a) calculate two-electron integrals and diagonalize the Fock matrix and (b) calculate a new density matrix until the convergence criterion is met. This process constitutes SCF. Of course, no one can guarantee the convergence (as we have already discussed in the previous section), on the other hand, since the variational principle is employed, the upper bound to the total energy is always obtained.

To really appreciate, what we have constructed here, let us mention a few interesting tricks or problems that are present during calculations in modern ab-initio packages:

- The initial guess is a crucial part of the whole process. It determines whether and how quickly the procedure will converge to a physically meaningful solution. To mention at least one method, Superposition of Atomic Densities builds the initial density as the spin-averaged sum of atomic Unrestricted Hartree-Fock computations in the current basis. When the geometry optimization task is being performed, results from previous SCF calculations may be reused as initial guess for the following ones, which obviously leads to significant speedup.

- One bottleneck is the calculation of two electron integrals. Most of them are zero or almost zero. Therefore, some guesses to predict which ones contribute to the result the most can be employed and subsequently only those which are expected to give larger contributions than the chosen threshold may be calculated.

- Most of the orbitals are unoccupied (so called virtual). During matrix inversions, we may want to use some simplification similar to the above-mentioned omitting of the negligible two electron integrals. Fortunately, several iterative approaches that consider only the subspace of occupied orbitals were developed – direct inversion of the iterative subspace [50] or its modern, parallel, order of magnitude faster and diagonalization-free alternative Q-next [51].

- Density matrices obtained during calculations frequently exhibit oscillatory behavior. This can be addressed by obtaining Fock matrix without calculating it from the density matrix, but by predicting it (fitting, neural network inference, etc.) from several (usually three) previously calculated matrices. This problem is crucial, since oscillations together with numerical errors arising from previously discussed issues might even lead to divergence [47].

Let us very briefly return to the treatment of spins introduced before equation (2.24). In restricted Hartree-Fock theory, a single molecular orbital is utilized twice, once multiplied by the α spin function and once by the β spin function

within the Slater determinant. Conversely, unrestricted Hartree-Fock theory employs distinct molecular orbitals for α and β electrons. This approach, also known as different orbitals for different spins [47], leads to two coupled Roothan equations

$$\begin{aligned} \mathbf{F}^{\alpha}\mathbf{C}^{\alpha} &= \mathbf{S}^{\alpha}\mathbf{C}^{\alpha}\boldsymbol{\varepsilon}^{\alpha}\,, \\ \mathbf{F}^{\beta}\mathbf{C}^{\beta} &= \mathbf{S}^{\beta}\mathbf{C}^{\beta}\boldsymbol{\varepsilon}^{\beta}\,. \end{aligned} \tag{2.31}$$

On the other hand, restricted open shell Hartree-Fock theory employs doubly occupied molecular orbitals whenever feasible and resorts to singly occupied orbitals for unpaired electrons only when necessary. This is probably the simplest model of open-shell molecules, yet it is extremely challenging to implement since restricted open shell Hartree-Fock theory leads to equation (2.30), where the Fock matrix $\mathbf{F}$ is unfortunately not determined uniquely.

2.3 Post Hartree-Fock methods

The Hartree-Fock method can describe the total energy of electronic systems with reasonable precission – Hartree-Fock method includes electrostatic interaction, Pauli repulsion and polarizability. The approximations used by Hartree-Fock are not precise enough for many chemical and physical processes and properties. From a numerical perspective, the difference between total energies is what is relevant for many processes. This difference is of an order of magnitude smaller than the total energies, and the errors in Hartree-Fock total energies don't have to cancel out when the energies are subtracted.

In general, the difference between the exact ground state and the ground state obtained from Hartree-Fock methods is called correlation – because of the employed variational principle in the SCF procedure, correlations have to lower the total energy obtained from the mean-field calculations. We may distinguish between two types of correlations – static and dynamic (yet this distinction is not well-defined in mathematical terms) [47].

Static correlation occurs when several states in the ground state wave function expansion of Slater determinant share equal significance or the groundstate is degenerate [52]. This makes the single determinant ansatz bad approximation for the electronic structure. The static correlation can be well addressed by Multi configurational SCF [47], where the ansatz comprises several Slater determinants

$$|\Phi(\{c_i\})\rangle = \sum_i c_i \,|\Psi_i\rangle \ . \tag{2.32}$$

During calculation, the coefficients c_i have to be found. The main obstacle of this approach is the selection of the set $\{|\Psi_i\rangle\}_i$. There exists several strategies how to do so [47, 52]; however, we will not devote more space to this issue, since it is not so relevant for molecular solids studied in this text.

Dynamic correlation is linked to the actual electron-electron repulsion and electron-nuclear attraction caused by the Coulomb force [52] in both short- and long-range regions. The most significant outcome of dynamic correlation in the calculation of molecular solids is associated with the slow convergence of basis

set extrapolations (see section (2.5)), arising from the behavior of electrons when they are in proximity.

When two electrons (or the electron and nuclei) coincide in space, the electronic Hamiltonian (2.2) becomes singular in the terms introducing Coulomb interaction. This leads to a non-differentiable condition for wave function also called Kato cusp condition[4]

$$\left.\frac{\partial\left(\frac{\Psi(r_{12})}{\Psi(0)}\right)}{\partial r_{12}}\right|_{r_{12}=0} = \text{const.} \tag{2.33}$$

A wave function in the form of a single Slater determinant cannot meet the non-differentiable requirement (2.33). Another issue, which is essential for molecular solids, is the dispersion interaction – resulting from long-range electron correlations, as was already described in the introduction. Dispersion interaction is on the level of SCF not included at all.

To incorporate the condition (2.33) into calculations, to expand the ansatz, and to include dispersion interactions, so-called excited Slater determinants $\Psi_{ij\ldots l}^{ab\ldots d}$ are considered. Using the creation and annihilation operators from the second quantization, one may write

$$|\Psi_{ij\ldots l}^{ab\ldots d}\rangle = \hat{c}_a^\dagger \hat{c}_b^\dagger \ldots \hat{c}_d^\dagger \hat{c}_i \hat{c}_j \ldots \hat{c}_l \, |\Psi_0\rangle \ . \tag{2.34}$$

Indexes $a,\ b,\ c,\ \ldots$ correspond to excitations and $i,\ j,\ k,\ \ldots$ to holes. Throughout this text, this convention is consistently applied. Excited Slater determinants are referred to as Slater determinants that are singly, doubly, triply, quadruply, etc., excited – these determinants are commonly denoted as Singles (S), Doubles (D), Triples (T), Quadruples (Q), and so forth.

The condition (2.33) is not the end of the story, since more strict assumptions might be made for the ansatz wave function. They are discussed in section (2.6) and the efforts to made asatzes compliant with them resulted in the theory of explicitly correlated methods, which are of extreme importance for molecular solids and precise energy calculations in general.

Configuration Interaction

CI uses variational method on wave functions of the form [47]

$$|\Psi\rangle = t_0 \, |\Psi_0\rangle + \sum_{ai} t_i^a \, |\Psi_i^a\rangle + \sum_{abij} t_{ij}^{ab} \, |\Psi_{ij}^{ab}\rangle + \ldots \tag{2.35}$$

with coefficients t_0,t_i^a, t_{ij}^{ab} and so on determined from the minimization condition. The order of the theory is determined by which excited determinants are considered (CIS – singles; CISD – singles and doubles; etc.) When all possible excitations are included, we use Full CI. The Hilbert space size of Full CI is given by all the possibilities to distribute n electrons in $2N$ orbitals (determined by the

[4]For simplicity, presented for the system of two electrons, where the only meaningful parameter is their distance r_{12}. The equation (2.33) was obtained by simplification of equation (1) presented in [53]. $\Psi(0)$ has to be obtained as an average on small sphere around the point zero.

number of basis functions), which is $\frac{(2N)!}{n!(2N-n!)}$ [47]. For specific systems, the size of the problem can be reduced by symmetry considerations; yet, the unfavorable scaling is still present. However, within system sizes where achieving the Full CI solution is feasible, it offers a highly valuable benchmark. The objective of nearly every computational approach is to achieve the closest approximation to the Full CI solution. This leads to statements where physicists claim that a method is exact to some order, meaning that all terms up to this order are present when compared with Full CI. In equation (2.35), ansatz of the form (2.32) may be introduced, leading to Multi-reference CI.

Møller-Plesset perturbation theory

An alternative way to the variational principle is to obtain the ground state as a perturbation of the solution calculated by the SCF procedure. Using the standard results from stationary perturbation theory with nondegenerate spectrum[5], Second order Møller-Pleset perturbation theory (MP2) can be obtained. The zero and first order energies together comprise the Hartree-Fock energy and hence, there are several ways how to partition these two lowest order contributions – needless to say, all possible ways lead to identical second and higher order perturbations and thus to identical contributions to correlation energy beyond the Hartree-Fock model.

The canonical way is to take the unperturbed Hamiltonian $\hat{H}_0$ as a sum over Fock operators (2.20). Since the summation of Fock operators results in a duplication of the average electron-electron repulsion [47], the perturbation is obtained as (in this case, it is not guaranteed that the perturbation is small)

$$\hat{V} = \underbrace{\sum_{i}^{N}\sum_{j>i}^{N}\hat{g}_{ij}}_{\text{exact}} - \underbrace{\sum_{i}^{N}\sum_{j}^{N}\langle\hat{g}_{ij}\rangle}_{\text{mean-field operator}} \quad . \tag{2.36}$$

By applying this perturbation, the wave function at zeroth order corresponds to the Hartree-Fock determinant, while the zeroth-order energy is simply the total sum of orbital energies. Originaly, derivation of this theory was done in a way to make the zeroth order energy equal to the total SCF energy and the first order correction then vanishes. See [49] for full derivation obtained from this condition.

[5]Meaning solving the problem

$$(\hat{H}_0 + \lambda\hat{V})\,|\Psi_j\rangle = E_j\,|\psi_j\rangle \ ,$$

where the solution for the $\hat{H}_0\,|\Psi_j^{(0)}\rangle = E_j^{(0)}\,|\Psi_j^{(0)}\rangle$ is known, in terms of power series

$$|\Psi_j\rangle = \sum_{i=0}^{\infty}\lambda^i\,|\Psi_j^{(i)}\rangle \ ,$$
$$E_j = \sum_{i=0}^{\infty}\lambda^i E_j^{(i)} \ ,$$

with additional constraint $\langle\Psi_j^{(0)}|\Psi_j^{(i)}\rangle = 0$ for $i \in \mathbb{N}$. The most important result is $E_j^{(i)} = \langle\Psi_j^{(0)}|\hat{V}|\Psi_j^{(i-1)}\rangle$.

In both cases, the contribution in the second order is described by equation ($\hat{V}$ is as in (2.36))

$$\langle \Psi_{ij}^{ab} | \, \hat{V} \, | \Psi_0 \rangle = (\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b) t_{ij}^{ab} \, . \tag{2.37}$$

From this equation, restricted amplitudes t_{ij}^{ab} can be straightforwardly obtained by application of the operators $\hat{g}_{ij}$ from equation (2.36). This leads to the second order energy correction

$$E^{\mathrm{MP2}} = \frac{1}{4} \sum_{ij,ab} t_{ij}^{ab} \, \langle ij||ab \rangle = \frac{1}{4} \sum_{ij,ab} \frac{|\langle ij||ab \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \, , \tag{2.38}$$

where we employ the spin orbital formalism, the definition $\langle ij||ab \rangle := \langle ij|ab \rangle - \langle ji|ab \rangle$ and the two-electron repulsion integral has the form [47] (note that chemists sometimes use a slightly different definition [47])

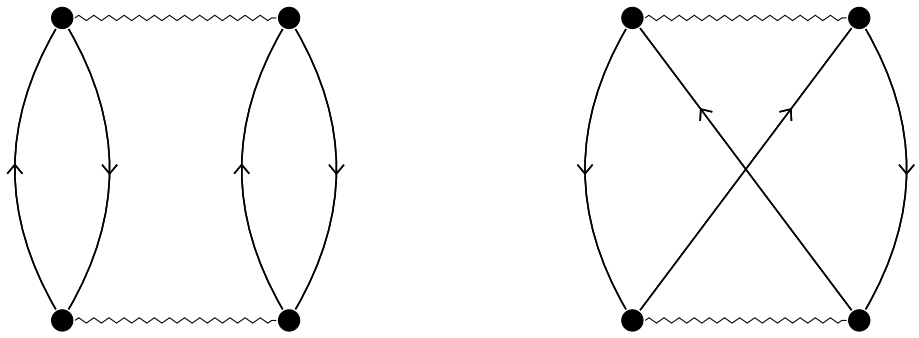
$$\langle ij|ab \rangle = \int \int \psi_i^*(\boldsymbol{r}_1) \psi_j^*(\boldsymbol{r}_2) \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \psi_a(\boldsymbol{r}_1) \psi_b(\boldsymbol{r}_2) \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \, . \tag{2.39}$$

Singly excited determinants do not contribute to the second order perturbation theory (but they do contribute to higher order theories) since the SCF orbitals are eigenfunctions of the Fock operator and $\langle \psi_0 | \hat{H} | \psi_a^i \rangle = 0$ is off-diagonal term of Fock matrix which was diagonalized by SCF calculation [47]. The equation $\langle \psi_0 | \hat{H} | \psi_a^i \rangle = 0$ is basically a statement of the so-called Brillouin theorem, which will be later also used for simplification of CC energy expressions. The equation (2.38) can be generalized into the form [49]

$$E^{\mathrm{MP2}} = \sum_{ij,ab} \frac{\langle ij|ab \rangle \left(2 \langle ab|ij \rangle - \langle ba|ij \rangle \right)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \, , \tag{2.40}$$

which can be displayed by Goldstone diagrams – see Figure (2.1).

Even though even fifth order implementation of the MP does exist, high orders of this method are rarely used since for the high computational cost associated with them the results are not as reliable as the ones obtained by other methods. Typical behavior of higher order MP calculation is shown in Figure (2.2). The MP methods can be optimized in many ways, and it can then describe systems of considerable sizes [54]. Some approaches are presented in the section (2.5).



(a) (b)

Figure 2.1: Goldstone diagrams can be used to depict contributions into energy in quantum chemistry. Here, we present direct term (a) corresponding to $\sum_{ij,ab} \frac{2|\langle ij|ab \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$ and exchange-like term (b) corresponding to $-\sum_{ij,ab} \frac{\langle ij|ab \rangle \langle ba|ij \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$ in equation (2.40).

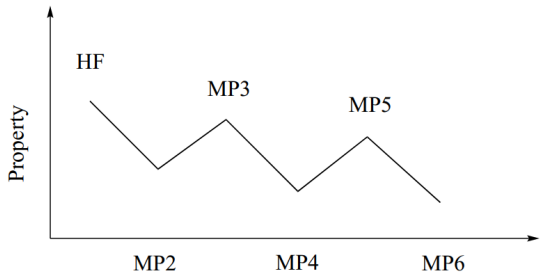


Figure 2.2: Oscillatory behavior of Finite order Møller-Plesset perturbation theory (MPn) (taken from [54]). MPn is known to "overcorrect" previous step, which leads to this typical zig-zag behavior.

Coupled clusters

CC was first introduced for the description of excitations in theory of nuclei [55, 56] and a decade later was modified [57, 58] for problems in chemical physics and many body physics in general[6]. To build the CC theory, we will follow [60]. First, we will define a cluster operator [60]

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots \hat{T}_N \,, \tag{2.41}$$

which is composed of a series of excitation operators T_i which introduce single Ψ_i^a, double Ψ_{ij}^{ab}, triple Ψ_{ijk}^{abc}, etc. excitations. The cluster operator (2.41) can be then used to obtain the wave function ansatz [60]

$$|\Psi\rangle = e^{\hat{T}} |\Psi_0\rangle = \left(\hat{I} + \hat{T} + \frac{1}{2}\hat{T}^2 + \frac{1}{6}\hat{T}^3 + \dots\right) |\Psi_0\rangle \tag{2.42}$$

where the operator $e^{\hat{T}}$ is sometimes called the wave operator[7]. The general structure of the wave operator is [60]

$$\begin{aligned} e^{\hat{T}} = \hat{1} + \underbrace{\hat{T}_1}_{\hat{C}_1} + \underbrace{\hat{T}_2 + \frac{1}{2}\hat{T}_1^2}_{\hat{C}_2} + \underbrace{\hat{T}_3 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3}_{\hat{C}_3} \\ + \underbrace{\hat{T}_4 + \frac{1}{2}\hat{T}_2^2 + \hat{T}_1\hat{T}_3 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 + \frac{1}{4!}\hat{T}_1^4}_{C_4} + \sum_{i>4} \hat{C}_i \,, \end{aligned} \tag{2.43}$$

where $\hat{C}_i$ corresponds to excitation in CI. We see that, for example, the operator $\hat{T}_2$ contributes even to higher excitations than just to $\hat{C}_2$ when compared with CI

[6] From the historical perspective, CC had initially much bigger success as quantum chemistry method rather than method in nuclear physics. Applications in nuclear physics followed only after using more powerful computers and advancing CC theory [59]. CC theories are currently hot topic in nuclear structure calculations.

[7] This is not the only way to go, we can for example replace the operator $e^{\hat{T}}$ in equation (2.42) with $e^{\hat{T}-\hat{T}^+}$, leading to unitary CC [59]. Or we can possibly reformulate the whole problem with a variational approach in mind and minimize $\frac{\langle\Psi_0|e^{\hat{T}^\dagger}\hat{H}e^{\hat{T}}|\Psi_0\rangle}{\langle\Psi_0|e^{\hat{T}^\dagger}e^{\hat{T}}|\Psi_0\rangle}$ which by the Rayleigh-Ritz variational principle should give us the upper bound for energy (see [61]). Unitary CC are sometimes used with advantage for electronic structure calculations; however, the variational approach is unpractical and requires further approximations [47]. We will stick with "classical" CC during our calculations.

theory and the operator $\hat{T}_2^2$ is thought to be more important than $\hat{T}_4$ for instance. In general, we can write (see equation (2.34) for comparison)

$$\hat{T}_n = (n!)^{-2} \sum t_{ij\ldots}^{ab\cdots} \hat{c}_a^\dagger \hat{c}_b^\dagger \ldots \hat{c}_i \hat{c}_j \,, \tag{2.44}$$

or when written explicitly using the notation with excited Slater determinants $\Psi_{ij\ldots}^{ab\cdots}$, we get [60]

$$\begin{aligned} \hat{T}_1 \Psi_0 &= \sum_{i,a} t_i^a \Psi_i^a \,, \\ \hat{T}_2 \Psi_0 &= \sum_{i>j,a>b} t_{ij}^{ab} \Psi_{ij}^{ab} \\ \hat{T}_3 \Psi_0 &= \sum_{i>j>k,a>b>c} t_{ijk}^{abc} \Psi_{ijk}^{abc} \end{aligned} \tag{2.45}$$

and so on.

The particular CC method is defined by the truncation of the cluster operator (2.41) which, of course, removes terms from the wave operator (2.43). The simplest imaginable approach is truncation after the $\hat{T}_1$ operator. However, this defines theory $\hat{T} = \hat{T}_1$ which does not in electronic structure theory contribute (the reasoning is the same as was in subsection (2.3)). Therefore, the simplest contribution arises from the cluster operator of the form $\hat{T} = \hat{T}_2$. This choice referred to as Coupled Clusters full double excitations (CCD) is hardly ever used since the operator $\hat{T}_1$ is neglected and CCD scaling is basically the same as the scaling when both $\hat{T}_1$ and $\hat{T}_2$ are considered. The first practical method Coupled Clusters full single and double excitations (CCSD) is thus given by the second order perturbation, defined by the excitation operator $\hat{T} = \hat{T}_1 + \hat{T}_2$, which by virtue of equation (2.43) is exact to the second order.

To calculate the CC ground state and energy, we have to determine the set of all contributing amplitudes $t_{ij\ldots}^{ab\cdots}$. Thereafter, the energy can be obtained in a quite straightforward manner. Firstly, we insert (2.42) into Schrödinger equation (2.1) and multiply by $e^{-\hat{T}}$ from the left [47]

$$e^{-\hat{T}} \hat{H} e^{\hat{T}} \left|\Psi_0\right\rangle = E \left|\Psi_0\right\rangle \,. \tag{2.46}$$

Subsequently, we multiply equation (2.46) by the reference state $\langle\Psi_0|$

$$\langle\Psi_0| e^{-\hat{T}} \hat{H} e^{\hat{T}} |\Psi_0\rangle = E \,. \tag{2.47}$$

The similarity transformed Hamiltonian $e^{-\hat{T}} \hat{H} e^{\hat{T}}$ in (2.47) can be then rewritten using Baker-Campbell-Hausdorff formula

$$\begin{aligned} e^{-\hat{T}} \hat{H} e^{\hat{T}} = \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{6}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] \\ + \frac{1}{24}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] + \ldots \,. \end{aligned} \tag{2.48}$$

For simplicity, we will now write results for CCSD $\hat{T} = \hat{T}_1 + \hat{T}_2$, however formulas for higher orders can be obtained by straightforward generalization of the presented procedure. In the electronic structure theory, the Hamiltonian consists

only of one– and two– electron operators. That gives us important simplification[8] summarized by the formula $\langle\Psi_0|\hat{H}\hat{T}_i|\Psi_0\rangle = 0$ for $i > 2$. Also $\langle\Psi_0|\hat{H}\hat{T}_1|\Psi_0\rangle = 0$ [47] thanks to the Brillouin theorem (again, see subsection (2.3)). Therefore, what remains after plugging equation (2.48) into (2.47) is

$$E = E_0 + \langle\Psi_0|[\hat{H}, \hat{T}_2]|\Psi_0\rangle + \frac{1}{2}\langle\Psi_0|[[\hat{H}, \hat{T}_1], \hat{T}_1]]|\Psi_0\rangle \ . \tag{2.49}$$

This can be, after some manipulations, rewritten into the form (see [49] for method of obtaining (un)restricted amplitudes t_i^a and t_{ij}^{ab} within the CCSD theory) [47, 49]

$$E = E_0 + \sum_{ij,ab}\left(\frac{1}{4}t_{ij}^{ab} + \frac{1}{2}t_i^a t_j^b\right)\left(\langle ij|ab\rangle - \langle ij|ba\rangle\right) \ . \tag{2.50}$$

Results from the CCSD method applied to molecular solids (see table 4 in [62] for instance) indicate that this method is not accurate enough to be able to describe for this thesis relevant systems reliably. Coupled Clusters full singles, doubles and triples (CCSDT) gives almost identical results when compared with full CI for many systems, where full CI is achievable [47] – on the other hand, the computational costs of CCSDT usually prevent it from being useful even for mid-size systems. To obtain computational technique with computational cost closer to CCSD and precision competing with CCSDT we can, in general, use two approaches – iterative or non-iterative [63]. The former methods allow triple excitation effects in the CCSD ansatz, whereas the later use converged CCSD wave functions that are perturbed. The best non-iterative method (i.e., method that is the most in agreement with CCSDT) is called Coupled Clusters full single, double and perturbative triple excitations (CCSD(T)) and uses Fourth order Møller-Pleset perturbation theory (MP4) with the perturbation coefficients replaced by CCSD amplitudes. Both iterative and non-iterative approaches scale as $O(N^7)$, which clearly is more practical than the $O(N^8)$ scaling of CCSDT (see table (2.1)). CCSD(T) construction might seem counterintuitive and the theory cannot be inferred from Hartree-Fock ansatz followed by standard perturbation techniques (for example CCSD(T) contains coupling between singles and triples which arises only in the fifth-order perturbation theory) [63]. The paper [63] gives mathematical perspective using the Lowdin's partitioning technique why CCSD(T) usually works so well.

2.4 Random Phase Approximation

The reason to study RPA is its universality – it describes ionic, covalent, metallic, and van der Waals bonding relatively well [49] – which is a direct constract to for example MPn which can diverge in metallic systems [49]. Many ways to obtain RPA can be found in literature, however since all of them are quite tedious and technical, we do not aim to do the full derivation in this text. At this point, the most natural derivation of RPA equations would be the diagrammatic simplification of the CC method – the tradeoff, depending on polarizability of the system, will be discussed in the section (2.5) using diagramatic considerations. Another

[8] This can be easily seen in the second quantization formalism.

is to use many-body Green's function formalism. Both of these strategies rely on approximate truncation of terms, and both of them lead to self-consistent calculations. However, there is an alternative formulation in which the RPA method is used to obtain an expression in the form of an integral over the frequency domain of a closed form expression for correlation energy using so-called linear response theory. The equation for correlation energy as well as the derivation can be found for example in [49] and [64]. The derivation is started by discussion of so-called adiabatic-connection, where the electrostatic interaction between electrons is slowly switched on by an external parameter α. The reference system without interaction ($\alpha = 0$) is usually called the Kohn-Sham system, and the general system is commonly referred to as α-interacting. A fully interacting system is given by $\alpha = 1$.

Thereafter, the fluctuation-dissipation theorem is applied. The fluctuation-dissipation theorem states that external perturbation influences the system in the same manner as spontaneous fluctuation. In mathematical terms, the fluctuation-dissipation theorem relates the density fluctuation operator to the density-density response function [49]. The correlation energy can be then evaluated by integration according to coupling constant α of terms depending on the density-density response function of the system described by the aforementioned parameter. All of these (now rather abstract) steps are exact, furthermore, there exists an exact expression for the density-density response function of the Kohn-Sham system χ_0. Unfortunately, approximation has to be made for the density-density response function of the α-interacting system χ_α. In this point, one might choose the density-density response function of α-interacting system to be equal to the fully interacting system, which defines the so called full-potential approximation. Choosing the χ_α to be constant is the simplest possible approach, yet its precision is not satisfactory. The more general (but more complicated) situation is to calculate χ_α as [49]

$$\chi_\alpha^{\mathrm{RPA}\,-1}(\omega) = \chi_0^{-1}(\omega) - \alpha\nu \tag{2.51}$$

which is much closer to the precise treatment [49]

$$\begin{aligned}\chi_\alpha(\mathbf{r},\mathbf{r}',\omega) =& \chi_0(\mathbf{r},\mathbf{r}',\omega) + \int \mathrm{d}^3\mathbf{r}'' \int \mathrm{d}^3\mathbf{r}''' \chi_0(\mathbf{r},\mathbf{r}'',\omega) \\ & \left(\alpha\nu(\mathbf{r}'',\mathbf{r}''',\omega) + f_{xc}^\alpha(\mathbf{r}'',\mathbf{r}''')\right) \chi_\alpha(\mathbf{r}''',\mathbf{r}',\omega)\,.\end{aligned} \tag{2.52}$$

where the unknown term f_{xc}^α is problematic and simply ignored in this case and $\nu(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|}$. This choice is referred to as RPA and it is the only approximation that is present in the derivation using the linear response theory. In practice, χ_α would not be calculated by equation (2.51) but rather by the standard treatment (meaning creation of the infinite serie by incorporating the left-hand side of the equation into the right-hand side) of the Dyson like equation (2.52) with $f_{xc}^\alpha = 0$. By algebraic manipulations (see [64]) the integral [49, 64]

$$E^{\mathrm{RPA}} = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \mathrm{Tr}\left(\ln\left(1 + \chi_0\nu\right) - \chi_0\nu\right) \tag{2.53}$$

is obtained, which is called the correlation within RPA theory. This is not completely consistent with our definition of correlation energy as the difference between SCF energy and the actual value, since the integral (2.53) also incorporates the electrostatic electron-electron interaction that is present in SCF calculations.

The integral (2.53) can replace problematic explicit correlation functionals in DFT leading, for example, to more accurate calculations within DFT with periodic boundary conditions, as was already noted in the introduction [25].

Another element of the RPA theory is the Casida equation (2.56). Within the linear response theory, it constitutes rewriting equation (2.51) by utilizing the summation over residues approach to integration in the complex plane [49]. However, the derivation by this procedure is again rather long, therefore, let us change strategy one more time. The advantage of the following approach is that it directly incorporates excitations, which are in time-dependent many-body theories described well by the Casida equation [65]. Let us therefore define the ground state to be an equation $\hat{Q}_\nu |\Psi\rangle = 0$, where $\hat{Q}^\dagger_\nu = \sum_{mi} X^\nu_{mi} \hat{a}^\dagger_m \hat{a}_i - \sum_{mi} Y^\nu_{mi} \hat{a}^\dagger_i \hat{a}_m$ is the simplest possible version of the excitation operator with desired properties for further calculations. Subsequently, we may wish to study expectation value [66]

$$\langle\Psi| \left[\hat{\mathcal{O}}, \left[\hat{H}, \hat{Q}^\dagger_\nu\right]\right] |\Psi\rangle = (E_\nu - E_0) \langle\Psi| \left[\hat{\mathcal{O}}, \hat{Q}^\dagger_\nu\right] |\Psi\rangle \ , \tag{2.54}$$

where the (for now general) operator $\hat{\mathcal{O}}$ is later set to

$$\delta\hat{Q}_\nu = \sum_{mi} \delta X^{*\nu}_{mi} \hat{a}_i{}^\dagger \hat{a}_m - \sum_{mi} \delta Y^{*\nu}_{mi} \hat{a}^\dagger_m \hat{a}_i \, . \tag{2.55}$$

In this manner, the following equation is straightforwardly obtained (see [66])

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \end{pmatrix} = (E_\nu - E_0) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X_n \\ Y_n \end{pmatrix} . \tag{2.56}$$

$\hat{Q}^\dagger_\nu = \sum_{mi} X^\nu_{mi} \hat{a}^\dagger_m \hat{a}_i$ leads to a slightly modified version of 2.56 where B is zero. This is the so-called Tamm-Dankoff approximation. After plugging $\delta\hat{Q}_\nu = \sum_{mi} \delta X^{*\nu}_{mi} \hat{a}_i{}^\dagger \hat{a}_m$ as operator $\hat{\mathcal{O}}$ in equation (2.54) we see that the right-hand side is equal to zero. However, the left-hand side in general doesn't have to be identically zero (interaction terms in Hamiltonian lead to nonzero contributions [66]). The applicability of this theory was extensively studied for many chemical and nuclear systems, with mixed results [67, 68].

RPA exhibits so-called self-correlation problem, which means that, even in a system with only one electron occupying orbital i, the correlation energy

$$E^{\mathrm{RPA}} = \frac{1}{2} \sum_{ab} \langle ii|ab\rangle \, t^{ab}_{ii} \neq 0 \tag{2.57}$$

is nonzero [49] (this "correlation" energy composes also of electron-electron electrostatic interaction, which in the case of one electron system is also expected to be zero). MP2 and CCSD methods do not encounter self-correlation, since they include exchange-like terms (such as $\langle ij||ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle$ in equations (2.38) and (2.50)) in their energy expressions. In the RPA theory, the inclusion of exchange-like contributions is called second order screened exchange [49], which is beyond the scope of this text.

2.5 Basis sets and scaling of methods

To be able to perform Hartree-Fock variational procedure and post Hartree Fock calculations, we have to specify the basis set in the equation (2.25) – i.e., spec-

ify the ansatz $\mathcal{A}$. We use so called correlation-consistent basis set [69, 70] cc-pVnZ (sometimes denoted in this text by VnZ) where the cardinal number n $\in \{\mathrm{D},\ \mathrm{T},\ \mathrm{Q},\ 5, \dots\}$ where D stands for double, T for tripple, etc. Correlation-consistent basis sets can be extended with diffusion functions to obtain so-called aug-cc-pVnZ (sometimes denoted for simplicity by AVnZ) which are, in general, superior for description of correlation. The binding energy can be obtained by the supermolecular approach as

$$E_{\text{intermolecular}} = E_{\text{tot}} - \sum_{i=1}^{N} E^{i}_{\text{mono}}\,, \tag{2.58}$$

where E_{tot} is the total energy of the whole system and E^{i}_{mono} is the energy of i-th monomer. There are two issues arising from the employed basis sets, which need to be addressed:

1. The theoretical framework introduced so far is in compliance with complete basis sets, however (aug)-cc-pVnZ is incomplete. Therefore, some kind of correction which would take into account the convergence of energy with respect to the size of basis sets is required. This procedure is not straightforward for the intermolecular interaction energy, since Hartree-Fock energy converges exponentially whereas correlations polynomialy. We use the technique of Helgaker et al. [71] which reads[9]

$$E^{\text{HF}}_{n} = E^{\text{HF}}_{\infty} + Ae^{-Bn}\,, \tag{2.61}$$

$$E_{n} = E_{\infty} + \frac{C}{n^{3}}\,, \tag{2.62}$$

   where E_n is the energy obtained with non-complete basis set (again n $\in\{$D, T, Q, 5, $\dots\}$) and E_∞ is the estimated value of energy evaluated in a complete basis set (see Figure (2.3)). This so-called Basis Set Incompleteness Error (BSIE) is more relevant for correlation contributions than SCF energy. The speed of convergence in equation (2.62) might not be suitable when the size of the system prevents usage of bigger basis sets. For calculations of molecular solids, this might be resolved by so-called explicitly correlated methods presented in the next section.

   Needless to say, convergence of the ground state energy does not necessarily imply the convergence to the real ground state of the system – the process which leads to converged energy (we may check this in plot for example) doesn't have to yield convergence of other physical/chemical properties or geometric properties after optimization.

[9] Better methods were proposed. For example, the correlation energy might be separated into singlet and triplet states and the correction might be performed by the set of equations [47]

$$E^{\text{singlet}}_{n} = E_{\infty} + \frac{C}{(n+\frac{1}{2})^{3}}\,, \tag{2.59}$$

$$E^{\text{triplet}}_{n} = E_{\infty} + \frac{C}{(n+\frac{1}{2})^{5}}\,, \tag{2.60}$$

which is connected to the different structures of singlet and triplet states when electrons are in proximity – see equations (2.71).

2. Basis Set Superposition Error (BSSE) arises from evaluating monomers in a smaller basis set than the whole system. This leads to a higher energy of monomers, which then overestimates binding. To mitigate this problem, one can employ the counterpoise correction method, which involves evaluating the energies of the monomers using the basis set of the whole system (see again Figure (2.3)).

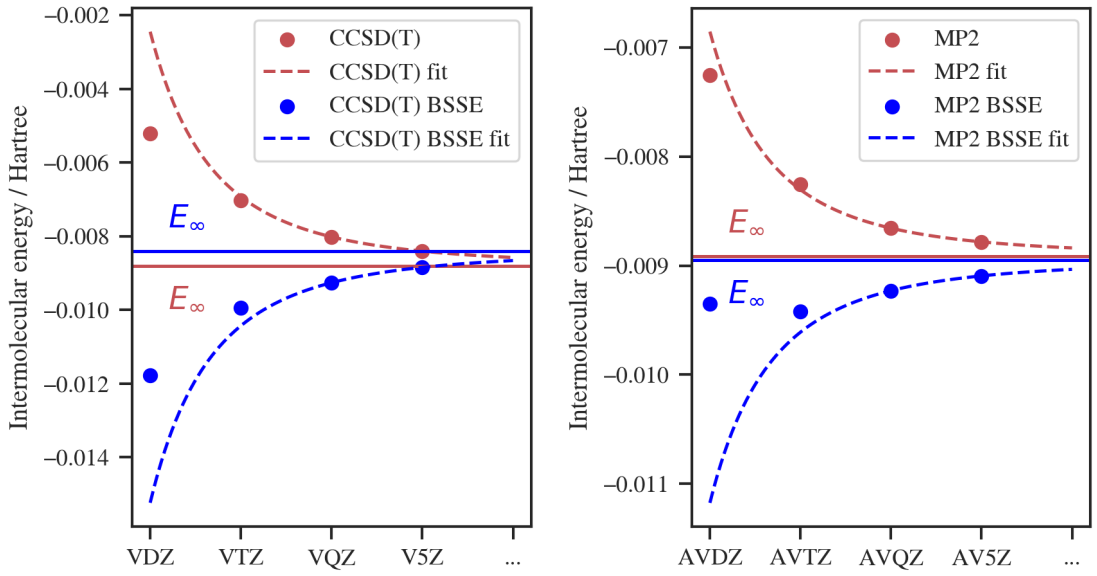
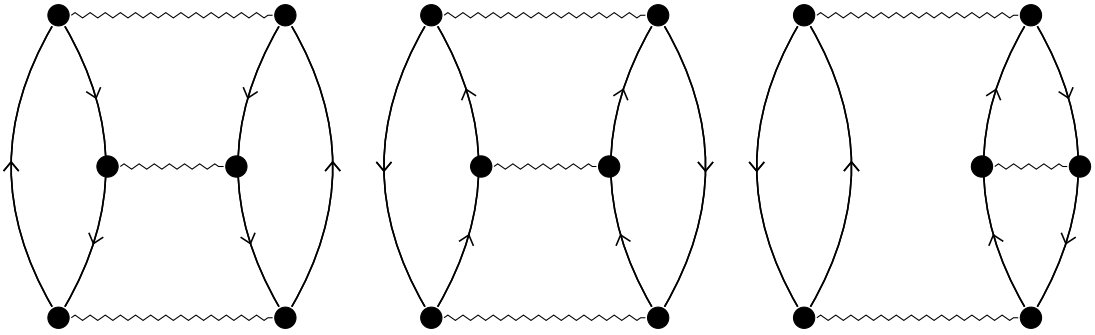


Figure 2.3: Convergence of intermolecular energy of hydrogen bonded methanol molecules with respect to the cardinal number of Dunning correlation-consistent basis set using the correction (2.62) calculated from the two largest basis sets. The intermolecular energy is much smaller than the total energy of the system, which is approximately -231 Hartree (MP2 with aug-cc-pVQZ). Furthermore, the MP2 and CCSD(T) intermolecular energy is entirely different from the SCF calculation, which gives approximately -0.0041 Hartree after the correction (2.61) calculated from aug-cc-pV5Z and aug-cc-pVQZ basis.

Table 2.1: Formal scaling of discussed computational methods [47, 63, 49]. N is describing the size of the problem. A more accurate description of scaling may be written as $O(N^k_{\text{occupied}} N^l_{\text{virtual}} N^m_{\text{basis}})$; furthermore, the computation might comprise more steps and scaling for only the most demanding one is shown.

Computational scheme	Scaling
RHF	$O(N^4)$
MP2	$O(N^5)$
CCD,CCSD, MP3	$O(N^6)$
CCSD(T), MP4	$O(N^7)$
CCSDT, MP5	$O(N^8)$
CCSDT(Q)	$O(N^9)$
CCSDTQ	$O(N^{10})$



(a) (b) (c)

Figure 2.4: Goldstone diagrams containing (a) particle-particle, (b) hole-hole and (c) particle-hole interaction.

The diagrams (such as the one in Figure (2.1) and (2.4)) that are taken into account in MP2, CCSD, and RPA allow us to evaluate the accuracy of these methods regarding polarizability of considered systems [49]. Particle-particle and hole-hole ladder diagrams, are regarded as relevant for weakly polarizable systems. Conversely, for strongly polarizable molecules, summation of all bubble diagrams and higher-order diagrams in general, should be significant. Particle-particle, hole-hole and hole-particle diagrams are present in CCSD together with infinitely many bubble diagrams. MP2 disregards particle-particle, particle-hole, and hole-hole ladder diagrams [49]. Additionally, MP2 incorporates solely a second-order bubble diagrams. Consequently, MP2 is anticipated to be precise exclusively for systems exhibiting intermediate polarizability. In the very similar manner as CCSD, RPA sums bubble diagrams up to infinite order (which makes it more reliable for highly polarizable systems than MPn methods); however, RPA is exact only to the first order [49].

Simplified scaling of computational methods based on the system size is in table (2.1). It is important to remember that CC, RPA, and MP2 energies are not variational (whereas SCF and CI are), therefore, no general claims on bounds of exact energy based on some post-Hartree-Fock methods are justified. Furthermore, the energy obtained by the equation (2.58) also cannot be bounded by the variational assumption.

There are other properties worth mentioning. Size extensivity requires the quantum-chemical method to scale linearly with the number of electrons. Size consistency (or sometimes referred to as strict separability) is a property that guarantees the consistency of the energy behavior when interaction between the involved molecular systems is nullified (usually by distances) [47]. It is expected, that when non-interacting systems are evaluated together in one calculation and separately, we should obtain the same total energy (after summing results from separate calculations). This might seem like a basic property that will hold for all quantum-chemical methods. However, it does not have to be the case for RHF or CI (whereas the Full CI is size consistent). On the other hand, all discussed methods (RHF, MPn,CC, CI, and RPA) are size extensive. Unfortunately, size consistency cannot say anything about the quality of description of separation (i.e., dissociation processes). For example, RHF inadequately describes disso-

ciation curves, rendering all post-HF methods dependent on RHF as an initial estimate unusable in this specific scenario [47]. When inappropriate numerical implementation of the ab-initio method is employed, even size-consistent methods might behave in a size inconsistent manner [72].

The local nature of DFT leads to approximate inclusion of electronic correlations. Treatment of short-distance correlation in SCF and post-HF might be improved by so-called Pair Natural and Localized Orbitals (which also leads to better convergence in equation (2.62)). Localization of orbitals is part of the group of tricks that aim to transform the scaling in table 2.1 into a more favorable one (sometimes even linear). Other tricks worth listing are Laplace transformed methods (briefly introduced at the end of this section on the example of MP2) and multipole expansion of physical properties. Density fitting approximates $\langle ab|ij\rangle$ as $\int \rho_{ai}(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\rho_{bj}(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2$ where ρ_{ai} and ρ_{bj} are expanded in auxiliary basis set as $\sum_\nu d_\nu\phi_\nu$ with coefficients d_ν determined in a data-driven manner [47]. These techniques allowed the method MBIE-SOS-AO-LT-MP2[10] to be used on an RNA system comprising 1664 atoms and 19,182 basis functions [73].

Laplace transformed second order Møller Plesset perturbation theory

The Laplace transform of a function $f : t \to f(t)$ is another function $\mathcal{L}\{f\} : s \to \mathcal{L}\{f\}(s)$ defined by equation

$$\mathcal{L}\{f\}(s) = \int_0^\infty f(t)e^{-st}\mathrm{d}t\,. \tag{2.63}$$

In this subsection, we will use the obvious property

$$\mathcal{L}\{1\}(s) = \frac{1}{s}\,. \tag{2.64}$$

By using the spin orbital formalism and the notation $\langle\cdot||\cdot\rangle$ we take MP2 energy

$$E^{\mathrm{MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle|^2}{\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j}\,, \tag{2.65}$$

where we are relying on the definition in equation (2.39). Now, by using equation (2.64) we may write

$$\frac{1}{\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j} = \int_0^\infty e^{-(\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j)t}\mathrm{d}t \tag{2.66}$$

and

$$E^{\mathrm{MP2}} = -\frac{1}{4}\int_0^\infty\sum_{ijab}|\langle ij||ab\rangle|^2 e^{-(\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j)t}\mathrm{d}t\,. \tag{2.67}$$

The dependency on t within the integrand can be shifted into the orbitals, where we are basically performing evolution in imaginary time

$$\begin{aligned}\psi_i(t) &= \psi_i(0)e^{\frac{1}{2}\epsilon_i t}\,,\\ \psi_a(t) &= \psi_a(0)e^{-\frac{1}{2}\epsilon_a t}\,.\end{aligned} \tag{2.68}$$

[10]Meaning multipole-based integral estimates – spin-opposite scaled – atomic orbitals – Laplace transformed MP2.

Thus, the energy can be rewritten as

$$E^{\mathrm{MP2}} = \int_0^\infty e(t)\mathrm{d}t\,, \tag{2.69}$$

where

$$e(t) = -\frac{1}{4}\sum_{ijab} |\,\langle a(t)b(t)||i(t)j(t)\rangle\,|^2\,. \tag{2.70}$$

This technique was first introduced for MP2 in [74]. It generalizes well to MPn and other methods such as RPA [75] or MP2 within periodic boundary conditions [76]. The main virtue of equations (2.69) and (2.70) is that one may reduce CPU time by saving some precalculated integrals on disk, since repeated evaluation of the fraction in (2.66) is quite costly. One could potentially achieve scaling that is up to one order more favorable than in table (2.1) – the tradeoff is on the other hand larger utilized memory.

2.6 Explicitly correlated methods

Explicitly correlated methods are able to circumvent the slow convergence (2.62) of the electron correlation energy regarding the basis set size, which arises from the poor description of the electron-electron cusp (see also condition (2.33)). For example, the wave function for the first-order singlet (S=0) and triplet (S=1) states of two electrons exhibit the following structure (see [77] or more general equations (2)–(5) in [53])

$$\begin{aligned} |\Psi_{S=0}\rangle &= \left(1+\frac{r_{12}}{2}\right)|\Psi_0(r_{12}=0)\rangle + \mathcal{O}(r_{12}^2)\,,\\ |\Psi_{S=1}\rangle &= \left(1+\frac{r_{12}}{4}\right)|\Psi_0(r_{12}=0)\rangle + \mathcal{O}(r_{12}^2)\,, \end{aligned} \tag{2.71}$$

near coalescence (i.e., the point $r_{12}=0$). By expansion of the finite number of Slater determinants, this condition cannot be attained [52]. Similarly to the whole section (2.3), which resulted from the condition (2.33), we may study a huge amount of explicitly correlated methods. Let us therefore mention a few approaches very briefly, which will hopefully give us understanding of the most basic elements of these methods.

One can directly incorporate terms ensuring non-differentiable cusp behavior of the exact ground state wave function into the ansatz. The first explicitly correlated approach to address (2.71) was developed by Hylleraas, who has written [52]

$$|\Psi(r_1,r_2)\rangle = e^{-\xi(r_1+r_2)}\sum_{lmn}\left(r_1^l r_2^m + r_1^m r_2^l\right)(r_1-r_2)^n\,, \tag{2.72}$$

as a famous ansatz for the wave function of helium atom. To generalize his approach to larger systems, so-called correlators $\hat{C} = \prod_{j>i}\hat{f}(\mathbf{r}_i,\mathbf{r}_j)$ can be introduced and used to modify the ground state ansatz $|\Psi\rangle = \hat{C}\,|\Psi_0\rangle$. Unfortunately, this simple generalization of (2.72) scales steeply with system size, which prevents achieving the defined objective of this section [52].

Transcorrelated methods

The transcorrelated approach introduces a similarity transformed Hamiltonian $\hat{\tilde{H}} = \hat{C}^{-1}\hat{H}\hat{C}$, known as the transcorrelated Hamiltonian ($\hat{C}$ is defined in the same manner as in the previous paragraph). This approach should involve only three electron integrals. Currently, it is not so commonly used in practical ab-initio calculations, yet transcorrelated CC methods were recently developed, and they promise some advantages ("drastic improvement of BSSE") [78, 79] over F12 methods, which will be introduced in the next paragraph. Furthermore, paper [78] showed that the possibly non-hermitian nature of the operator $\hat{\tilde{H}}$ doesn't cause many numerical issues during calculations.

R12/F12-method

The R12 approach adds two-electron basis functions into the traditional Slater determinant. In the simplest possible case of only two electrons without considering any excited Slater determinants, the R12 wave function can be written in the form [52]

$$|\Psi\rangle = (1 + \lambda\hat{Q}_{12}r_{12})\,|\Psi_0\rangle\ , \tag{2.73}$$

which is basically trying to mimic the structure of (2.71). The parameters λ has to be found by optimization. $\hat{Q}_{12}$ is a so-called strong orthogonality projector. Its job is to make $r_{12}\,|\Psi_0\rangle$ orthogonal to $|\Psi_0\rangle$ (and possibly to excited Slatter determinants if included). By modifying the R12-method where r_{12} in equation (2.73) is replaced by $f_{12} = -\frac{e^{-\xi r_{12}}}{\xi}$ [77] the famous F12-method is introduced (similarly to equation (2.72), the parameters ξ has to be found). It is probably the most used explicitly correlated method in quantum chemistry today. F12 methods are usually implemented for singles and doubles; however, a problem arises with triples, where the calculation of explicitly correlated contributions is hard. Therefore, employing scaling of explicitly correlated triple contributions to energies based on explicitly correlated contributions to singles and doubles is sometimes needed [80]. In the field of molecular solids, this scaling usually gives meaningful contributions for dimers. For trimers, tetramers, ... this scaling gives unsatisfactory results [81].

3. Distributed multipoles

DMA was introduced by A. J. Stone to precisely describe the electrostatic effects of big molecules. DMA performs multipole expansion into spherical harmonics around so-called sites (usually corresponding to atoms). Hence, even the first order of expansion introduces dipoles, quadrupoles, etc. when compared with expansion around only one point on the molecule. In theory, DMA can be used to describe electrostatic effects (including torque) quite accurately, even for rather close molecules. Multipole methods can be used for efficient large scale biomolecular simulations [82] and can be, in some cases, predicted from the structure of molecules utilizing neural networks [83]. We also hope that its precision allows us to evaluate electrostatic contributions to the binding energies of dimers in several molecular solids.

Methods based on multipole expansion and their interactions were successfully used to describe systems driven by Laplace [84], Helmholtz [85], Maxwell [86] and Stokes [87] equations. The Fast Multipole Method can be used to describe long-range interactions by expanding an arbitrary green function of studied systems [88]. There even exists application of multipole expansion for EFT and linearized gravity, where the expansion is done at the level of the action [89]. The theory of multipole expansion in the spherical tensor formalism and their interactions is well established.

In this chapter, we will recapitulate basic results from the representation theory and angular momentum algebra, define the SO(3) group, Wigner D matrices, and spherical harmonics. We will continue by deriving the formula for calculating interaction between multipoles in spherical harmonics basis (using the reference [90]). Finally, we will show, how to rotate multipoles and how to move the center of the multipole expansion.

The alternative approach to multipole expansion into the spherical harmonic basis is the usage of expansion into Cartesian tensors. It can be well-used for systems with simpler geometry, and it is probably easier to use for extrapolation/interpolation/fitting of precalculated ab-initio interaction energies on simpler geometries up to low order of expansion. The Cartesian tensor formalism can also straightforwardly include polarization, induction, and dispersion effects. We, on the other hand, wish to evaluate multipoles directly from the density matrix obtained by ab-initio calculations for a general system without any assumptions on the geometry. Furthermore, the spherical tensor formalism allows elementary generalization of the code to include multipoles up to arbitrary order and a more efficient and transparent way to store the coefficients of the expansion (which makes fitting of those coefficients on general geometry much simpler). Straightforward implementation of rotations and translations of the origin of the multipole expansion in the spherical harmonics basis is also possible.

3.1 Flight over representation theory

In this section, we describe the three most basic definitions from the theory of group representations, which will be used in the discussion of SO(3) group as well as in derivation, that will give an explicit equation for multipoles in the spherical

tensor formalism and formula for their interactions. In this text, G is a group and V is a (finite dimensional) vector space. These definitions are compiled from the chapter 3 in [91].

Definition 1 (Representation of group)**.** If there exists a homomorphism from G to a group of operators $U(G)$ acting on V, then $U(G)$ is called a representation of group G. If the homomorphism is, in fact, an isomorphism, we call the representation faithful. If there exists another group representation $U'(G)$ and matrix S such that

$$SU'(g)S^{-1} = U(g) \tag{3.1}$$

for every $g \in G$, we say that representation $U(G)$ and $U'(G)$ are equivalent.

Definition 2 (Invariant subspace)**.** Let W be the subspace of V. If for every $g \in G$ and $x \in W$ $U(G)x \in W$ then we call W the invariant subspace.

Definition 3 (Reducible and Irreducible representation)**.** Let $W \subset V$ be an invariant subspace. Let $\mathcal{B}$ be an orthogonal basis of V such that first n vectors in $\mathcal{B}$ generate the space W. Then we can write

$$[U(g)]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} D^{(1)}(g) & X(g) \\ 0 & D^{(2)}(g) \end{pmatrix} \tag{3.2}$$

and we say that $U(G)$ is reducible representation. From the multiplication properties of matrices, we can see that for every $g_1, g_2 \in G$

$$D^{(i)}(g_1 g_2) = D^{(i)}(g_1) D^{(i)}(g_2)\,, \tag{3.3}$$

therefore $U^{(1)}(G)$ and $U^{(2)}(G)$ are also group representations.

If the orthogonal complement of W in V is also reducible, then the representation $U(G)$ is called fully reducible or decomposable and $X(g)$ in equation (3.2) is zero.

3.2 Some results from the angular momentum algebra

One motivation for the formalism of spherical tensors is the possibility to transform wave functions for particles of arbitrary spin. Therefore, let us mention the basic properties of angular momentum operators and states in quantum mechanics. Definitions are taken from [91, 92].

The group of all proper orthogonal rotations (without reflections) in three dimensions is called the Special Orthogonal group (SO(3))[1], which can be parametrized as [91]

$$\hat{U} = \exp\left(-i\phi \mathbf{n} \cdot \hat{\mathbf{L}}\right) \tag{3.4}$$

[1]From a mathematical perspective, an Orthogonal group is a group of $n \times n$ orthogonal matrices, where the group operation is given by matrix multiplication. The orthogonal group breaks into two connected components, where the Special Orthogonal group is the component with determinant equal to 1 – meaning that it contains the identity element [91].

where $\mathbf{n}$ is a unit vector in the direction of the axis of rotation and $\hat{\mathbf{L}} = (\hat{L}_1, \hat{L}_2, \hat{L}_3)$ is the angular momentum operator. Since the generators satisfy the commutation relations

$$[\hat{L}_1, \hat{L}_2] = i\hat{L}_3\,, \quad [\hat{L}_3, \hat{L}_1] = i\hat{L}_2 \quad , [\hat{L}_2, \hat{L}_3] = i\hat{L}_1\,, \tag{3.5}$$

the SO(3) is a Lie group (the equation (3.4) implies differentiability and smoothness). Usually, we construct a Casimir operator $\hat{L}^2 = \hat{L}_1^2 + \hat{L}_2^2 + \hat{L}_3^2$ (operator commuting with all generators) and then find basis as $|lm\rangle$ where

$$\hat{L}^2\,|lm\rangle = l(l+1)\,|lm\rangle\ , \quad \hat{L}_3\,|lm\rangle = m\,|lm\rangle\ . \tag{3.6}$$

After comparing the matrix elements $\langle lm|\hat{L}_i|l'm'\rangle$ with the equation (3.2) we see that this group representation has a form, where the states with different l are not mixed and the SO(3) group is divided into irreducible representations of the sizes of $2l+1$. The quantum number m of the generator $\hat{L}_z$ then labels the basis states within that irreducible representation, which are representations of the SO(2) group (thus this is a parametrization of rotations around the z axis) [91]

$$\mathrm{SO}(2) \subset \mathrm{SO}(3)\,. \tag{3.7}$$

To parametrize rotations in our code that calculates spherical harmonics and interaction of multipoles, we utilize quaternions, where we employ isomorphism between the group $SU(2)$ (Lie group of 2×2 unitary matrices with determinant 1) and quaternions with unit norm

$$\alpha + \beta\boldsymbol{j} = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} \longleftrightarrow \begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix} \tag{3.8}$$

and surjective homomorphism of $SU(2)$ onto $SO(3)$ [91].

In this text, α, β, and γ are Euler angles in the z–y–z convention in active interpretation (meaning that the rigid body is rotated, but the axes of the coordinate system are conserved). Furthermore, during the derivation of multipole interaction, we use the spherical angle Ω. The triples (r, θ, φ) denote spherical coordinates. The unit sphere is the set $\mathbb{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : ||x||_2 = 1\}$ where $||\cdot||_2$ is the quadratic norm.

Let us examine the vector space comprising the direct product of eigenstates associated with two angular momentum operators. The unitary transformation that connects the basis of uncoupled (product) states to those coupled to a definite total angular momentum is expressed as

$$|l_1 l_2 JM\rangle = \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} |l_1 m_1 l_2 m_2\rangle\, C_{l_1,l_2,J}^{m_1,m_2,M} \tag{3.9}$$

where the entity $C_{l_1,l_2,J}^{m_1,m_2,M}$ is called Clebsch-Gordan coefficient[2], which can be chosen to be real. They can also be expressed in terms of so-called Wigner 3j-symbols

$$C_{l_1,l_2,l_3}^{m_1,m_2,-m_3} = (-1)^{-l_1+l_2+m_3}\sqrt{2l_3+1}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\,. \tag{3.10}$$

[2]Today, they are usually written as C_{l_1,m_1,l_2,m_2}^{JM}; yet, we will use the same notation as in [90].

We continue this section by defining spherical harmonics[3]

$$Y_{lm}(\theta,\varphi) = (-1)^m \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos(\theta)) e^{im\varphi}\,, \tag{3.11}$$

which is the eigenstate of angular momentum operators $\hat{L}^2$ and $\hat{L}_z$. From the mathematical standpoint, spherical harmonics are basis functions for irreducible representations of SO(3). We will use Wigner D matrices

$$D^l_{m'm}(\alpha,\beta,\gamma) = \langle lm'|e^{-i\alpha\hat{L}_3}e^{-i\beta\hat{L}_2}e^{-i\gamma\hat{L}_3}|jm\rangle \tag{3.12}$$

to evaluate spherical harmonics, to evaluate electrostatic interaction and to rotate the multipole expansion. The Wigner D matrix is quite a complicated object. Since we have chosen the z direction to be rather special, it can be simplified into the form

$$D^j_{m'm}(\alpha,\beta,\gamma) = e^{-im'\alpha} d^j_{m'm}(\beta) e^{-im\gamma}\,, \tag{3.13}$$

where $d^j_{m'm}$ is a small Wigner d matrix (which is real in the z–y–z convention of Euler angles in active interpretation). This matrix will be (in different notation and convention) of considerable importance in the appendix (A).

Let us write without proof several useful identities that will be applied later on

$$Y_{lm}(\Omega) = (-1)^m Y^*_{l,-m}(\Omega)\,, \tag{3.14}$$

$$Y_{lm}(-\Omega) = (-)^l Y_{lm}(\Omega)\,, \tag{3.15}$$

$$D^l_{m0}(\alpha,\beta,\gamma) = (-1)^m \sqrt{\frac{4\pi}{2l+1}} Y_{l,-m}(\beta,\alpha)\,, \tag{3.16}$$

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \delta(m_1+m_2+m_3)(-1)^{l_1-l_2-m_3} K^{l_1,l_2,l_3}_{m_1,m_2,m_3}\,,\,, \tag{3.17}$$

$$Y_{lm}(0,\varphi) = \sqrt{\frac{2l+1}{4\pi}}\delta_{m0}\,, \tag{3.18}$$

where $K^{l_1,l_2,l_3}_{m_1,m_2,m_3}$ is positive and symmetrical regarding swapping indexes 1 and 2. The equation (3.14) will allow us to elegantly encode coefficients into real numbers, while equation (3.17) will reduce the number of iterations in the energy calculation.

As a direct consequence of Schur orthogonality theorem[4], we may write

$$\begin{aligned}\int_0^{2\pi}\mathrm{d}\alpha\int_0^{\pi}\mathrm{d}\beta\sin\beta\int_0^{2\pi}\mathrm{d}\gamma\, D^l_{mk}(\alpha,\beta,\gamma)^* D^{l'}_{m'k'}(\alpha,\beta,\gamma) = \\ = \frac{8\pi^2}{2l+1}\delta_{m'm}\delta_{k'k}\delta_{l'l}\,.\end{aligned} \tag{3.19}$$

[3]Associated Legendre polynomials are defined as $P_l^m = (1-x^2)^{\frac{m}{2}}\frac{\mathrm{d}^m}{\mathrm{d}x^m}(P_l(x))$ for $m \in \{0,\ldots,l\}$ and Legendre polynomials as $P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}\left((x^2-1)^l\right)$, i.e., the Condon–Shortley phase is included in the definition of spherical harmonics, not in the definition of associated Legendre polynomials, but it doesn't matter much.

[4]It states that the matrix elements of irreducible representations of a finite group form an orthonormal set under a certain inner product.

By using the relation (3.16), we may rewrite equation (3.19) as

$$\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\pi} \mathrm{d}\beta \, \sin\beta \, Y_{l'm'}\left(\beta,\alpha\right) Y^*_{lm}\left(\beta,\alpha\right) = \delta_{ll'}\delta_{mm'} \,. \tag{3.20}$$

Combining this result with the explicit formula $Y_{00} = \frac{1}{\sqrt{4\pi}}$ gives us a relation that will show its importance in the discussion of angular quadratures in section (3.5)

$$\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\pi} \mathrm{d}\beta \, \sin\beta \, Y_{lm}\left(\alpha,\beta\right) = \begin{cases} \sqrt{4\pi}\,, & \text{if } l = m = 0\,; \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3.21}$$

The orthogonality relations (3.20) together with equations (for any function f defined on unit sphere $\mathbb{S}^2$ and space $\mathbb{R}^3$ respectively)

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-1}^{l} c_{lm} Y_{lm}(\theta,\varphi) \quad \text{respectively,} \tag{3.22}$$

$$f(r,\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-1}^{l} c_{lm}(r) Y_{lm}(\theta,\varphi)\,. \tag{3.23}$$

mean that spherical harmonics form an orthogonal basis in the space of functions defined on the unit sphere $\mathbb{S}^2$.

3.3 Interaction of two multipoles in the spherical tensor formalism

The first-ever derivation of the interaction of two multipoles was presented in [93], which used group theoretical considerations to obtain interaction energy between two distributions of charges. Since then, several approaches have been proposed (some of them use group theory, some [94] properties of irreducible tensors). We will use the group theoretical approach presented in [90] which is, to our knowledge, the simplest one (yet the first step from the three-part derivation is not described well, and we had to change it a little).

Let us have two non-overlapping distributions of charge. To describe this situation in detailed geometrical terms, we note the point around which the first charge distribution is centered by O_1 and conversely the second one centered at O_2. Furthermore, let the vector $O_2 - O_1 = \boldsymbol{R}$ be described by distance R and spherical angle Ω_{12}. We can choose an arbitrary charge in the first distribution and describe its position relative to O_1 by r_1 and Ω_1 and the same thing can be done for the second distribution. The only condition for the geometry is that the two distributions do not overlap

$$|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1| > 0\,, \tag{3.24}$$

for all possible $\mathbf{r}_1$ and $\mathbf{r}_2$.

Now, we can write the $\frac{1}{r_{12}} = \frac{1}{|\mathbf{R}+\mathbf{r}_2-\mathbf{r}_1|}$ as a sum of products of spherical harmonics[5]

$$\frac{1}{r_{12}} = \sum_{l_1,l_2,l} \sum_{m_1,m_2,m} A^{m_1,m_2,m}_{1_2,l_2,l}(r_1,r_2,R) Y_{l_1m_1}(\Omega_1) Y_{l_2m_2}(\Omega_2) Y^*_{lm}(\Omega_{12})\,. \tag{3.25}$$

[5]By viewing the fraction $\frac{1}{r_{12}}$ as a function of $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{R}$ and repeated application of equation (3.23).

Thus, the goal of the derivation is to find an explicit form for the coefficient $A^{m_1,m_2,m}_{1_2,l_2,l}(r_1,r_2,R)$ and then to multiply the expression (3.25) by charge densities and integrate it over both distributions of charges to obtain the interaction energy. We will proceed in three steps:

1. **Expanding the coefficients $\mathbf{A^{m_1m_2m}_{l_1l_2l}}$ based on invariance conditions**.

   When viewing the term $\frac{1}{r_{12}}$ as a function of $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{R}$, it is invariant under simultaneous rotation of these vectors. Due to the discussed structure of the group SO(3), the subspace $l_1 \otimes l_2 \otimes l$, which is spanned by the products of three spherical harmonics corresponding to a definite value of (l_1, l_2, l), is reducible. The whole term

   $$\sum_{m_1,m_2,m} A^{m_1,m_2,m}_{1_2,l_2,l}(r_1,r_2,R) Y_{l_1m_1}(\Omega_1) Y_{l_2m_2}(\Omega_2) Y^*_{lm}(\Omega_{12}) \tag{3.26}$$

   therefore needs to transform as an identity (the reducibility of the subspace $l_1 \otimes l_2 \otimes l$ allowed us to consider all values (l_1, l_2, l) separately). Since the rotation transformation doesn't change magnitudes of vectors, we may split the term $A^{m_1,m_2,m}_{1_2,l_2,l}(r_1,r_2,R)$ as $A_{1_2,l_2,l}(r_1,r_2,R) B^{m_1,m_2,m}_{1_2,l_2,l}$ where $A_{1_2,l_2,l}(r_1,r_2,R)$ is independent of the rotation, and

   $$\sum_{m_1,m_2,m} B^{m_1,m_2,m}_{1_2,l_2,l} Y_{l_1m_1}(\Omega_1) Y_{l_2m_2}(\Omega_2) Y^*_{lm}(\Omega_{12})\,, \tag{3.27}$$

   has to transform as an identity. This can be achieved by setting $B^{m_1,m_2,m}_{1_1,l_2,l}$ equal to Clebsch-Gordan coefficient (see references in [90]).

   The rotation invariance together with the relation (3.15) means that l_1+l_2+l is even. We may therefore rewrite (3.25) as

   $$\begin{aligned} \frac{1}{r_{12}} = & \sum_{l_1,l_2,l; l_1+l_2+l \text{ is even}} A_{l_1,l_2,l}(r_1,r_2,R) \\ & \sum_{m_1,m_2,m} C^{m_1m_2m}_{l_1l_2l}\, Y_{l_1m_1}(\Omega_1)\, Y_{l_2m_2}(\Omega_2)\, Y^*_{lm}(\Omega)\,, \end{aligned} \tag{3.28}$$

   Another obvious property of the expression (3.25) is that both sides are real. Furthermore, the expression

   $$\sum_{m_1,m_2,m} C^{m_1m_2m}_{l_1l_2l}\, Y_{l_1m_1}(\Omega_1)\, Y_{l_2m_2}(\Omega_2)\, Y^*_{lm}(\Omega) \tag{3.29}$$

   in equation 3.28 is also real and therefore

   $$A_{l_1,l_2,l}(r_1,r_2,R) \in \mathbb{R}\,. \tag{3.30}$$

2. **Proving that non-vanishing terms have to follow relation $\mathbf{l = l_1 + l_2}$.**
   This will be done using the non overlapping condition (3.24) and the fact that $r_{12}^{-1} = |\mathbf{r}_2 + \mathbf{R} - \mathbf{r}_1|$ is simultaneously solution of the following Laplace equations

   $$\Delta_{r_1}\left(\frac{1}{r_{12}}\right) = 0\,, \;\; \Delta_{r_2}\left(\frac{1}{r_{12}}\right) = 0\,, \;\; \Delta_{R}\left(\frac{1}{r_{12}}\right) = 0\,, \tag{3.31}$$

where the index of Laplacian indicates in which variable the equation is solved. The solutions of Laplace equation are well known

$$r^l\, Y_{lm}\,, \quad r^{-l-1}\, Y_{lm}\,. \tag{3.32}$$

Combining this with the structure of (3.28), it is clear that one option how to rewrite the unknown coefficients in (3.28) is

$$A_{l_1,l_2,l}(r_1, r_2, R) = \frac{r_1^{l_1} r_2^{l_2}}{R^{l+1}} A_{l_1 l_2 l}\,. \tag{3.33}$$

Obviously, solution (3.33) satisfies the boundary condition for $\mathbf{R} \to \infty$ and the regularity condition for $\mathbf{r}_1 \to 0$ and for $\mathbf{r}_2 \to 0$. Due to these conditions, none of the seven remaining terms, which correspond to other combinations of solutions (3.32) of Laplace's equations (3.31), appear in the equation (3.33). We may also employ dimensionality considerations, leading to the aforementioned condition $l_1 + l_2 = l$. Thus, the expression (3.28) was simplified into the form

$$\begin{aligned}\frac{1}{r_{12}} =& \sum_{l_1,l_2} A_{l_1,l_2} \frac{r_1^{l_1} r_2^{l_2}}{R^{l_1+l_2+1}} \\ & \sum_{m_1,m_2,m} C_{l_1 l_2 l}^{m_1 m_2 m}\, Y_{l_1 m_1}(\Omega_1)\, Y_{l_2 m_2}(\Omega_2)\, Y_{lm}^*(\Omega)\,,\end{aligned} \tag{3.34}$$

where we employ the notation $A_{l_1,l_2} = A_{l_1,l_2,l_1+l_2}$.

3. **Finding the remaining coefficients $\mathbf{A_{l_1 l_2}}$ in (3.34).**

   The point of replacing $A_{l_1,l_2,l}(r_1, r_2, R)$ in equation (3.25) by A_{l_1,l_2} (3.34) was to get rid of the dependence of these coefficients on the geometry of the problem – for the coefficients A_{l_1,l_2} only the condition (3.24) remains relevant. Therefore, we consider the simplest possible case when $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{R}$ are parallel to the polar axis. In this simplified settings, it is easy to evaluate A_{l_1,l_2}.

   First, we will expand $|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|^{-1} = (R + r_2 - r_1)^{-1}$ using binomial expansion

$$\begin{aligned}\frac{1}{R + r_2 - r_1} &= \frac{1}{R}\left(1 - \frac{r_1 - r_2}{R}\right)^{-1} \\ &= \sum_{l_1 l_2} (-1)^{l_2} \frac{(l_1 + l_2)!}{l_1! l_2!} (r_1^{l_1} r_2^{l_2} / R^{l_1+l_2+1})\,.\end{aligned} \tag{3.35}$$

   Then, we use relation (3.18) to rewrite equation (3.34) for these simplified conditions

$$\begin{aligned}\frac{1}{R + r_2 - r_1} =& \sum_{l_1,l_2} A_{l_1,l_2} r_1^{l_1} r_2^{l_2} / R^{l_1+l_2+1} C_{l_1,l_2,l_1+l_2}^{0,0,0} \\ & \left(\frac{(2l_1+1)(2l_2+1)(2l_1+2l_2+1)}{(4\pi)^3}\right)\end{aligned} \tag{3.36}$$

and compare results using the known value [90, 95]

$$C^{000}_{l_1 l_2 l} = \left(\frac{l!}{l_1! l_2!}\right)\left(\frac{(2l_1)!(2l_2)!}{(2l)!}\right)^{\frac{1}{2}} . \tag{3.37}$$

This gives us (quite an ugly) expression

$$\begin{aligned}\frac{1}{r_{12}} =& 4\pi \sum_{l_1,l_2=0}^{\infty} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \sum_{m=-l_1-l_2}^{l_1+l_2} \frac{(-1)^{l_2}}{2l+1}\sqrt{\frac{4\pi(2l+1)!}{(2l_1+1)!(2l_2+1)!}} \\ & \frac{r_1^{l_1} r_2^{l_2}}{R^{l+1}} C^{m_1,m_2,m}_{l_1,l_2,l}\, Y_{l_1,m_1}(\Omega_1)\, Y_{l_2,m_2}(\Omega_2) Y^*_{l_1+l_2,m}(\Omega_{12}) .\end{aligned} \tag{3.38}$$

This expression has to be integrated over both distributions of charges. Therefore, we define

$$Q^{(i)}_{lm} = \int \rho^{(i)} \sqrt{\frac{4\pi}{2l+1}} |r^{(i)}|^l\, Y_{lm}(\Omega_i) \mathrm{d}\lambda^3 , \tag{3.39}$$

where $r^{(i)}$ is distance taken from O_i and $\rho^{(i)}$ is the charge density of i-th charge distribution. By using the definition

$$[l_1, l_2] = \sqrt{\frac{4\pi(2l)!}{(2l_1)!(2l_2)!}} = \sqrt{4\pi \frac{(2l_1+2l_2)!}{(2l_1)!(2l_2)!}} , \tag{3.40}$$

we may simplify the equation (3.38) with the help of Wigner 3j symbol into the final form

$$\begin{aligned} V = & \sum_{l_1,l_2=0}^{\infty} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \sum_{m=-l_1-l_2}^{l_1+l_2} [l_1, l_2] R^{-l_1-l_2-1} Q^{(1)}_{l_1,m_1} Q^{(2)}_{l_2,m_2} \\ & (-1)^{l_1} Y_{l_1+l_2,m}(\Omega_{12}) \begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix} .\end{aligned} \tag{3.41}$$

And that is it. Note, that the papers [96, 97] present different forms of equation (3.41), since they use different definitions. Furthermore, the implementation in our code calculating the energy had to be slightly modified, since the spherical package [98], which we use for the calculation of spherical harmonics (see appendix (A)), uses entirely different norms from the aforementioned papers. The implemented equation needed careful "balancing" of all definitions, checking phase conventions as well as different interpretations of Euler angles and quaternions.

3.4 Rotating and moving multipoles in the spherical harmonics basis

Rotation and movement of multipole expansion is regularly used in multipole methods [99]. In our case, the movement of the center of multipole expansion is used by code integrating density matrix and the rotation of multipole expansion

is helpful for reducing the number of calculations as well as for testing purposes. Let us therefore devote a few words to these techniques.

Since Wigner D matrices form a representation of the rotation group, it follows that

$$D^l_{mm'}(\Omega + \Delta\Omega) = \sum_{m''=-l}^{l} D^l_{mm''}(\Delta\Omega) D^l_{m''m'}(\Omega) \tag{3.42}$$

Using the relation (3.16) we can rewrite equation (3.42) into the form that rotates spherical harmonics

$$Y_{lm}(\Omega + \Delta\Omega) = \sum_{m'=-l}^{l} D^l_{mm'}(\Delta\Omega)^* Y_{lm'}(\Omega)\,. \tag{3.43}$$

Rewriting this in Q_{lk} terms (since Q_{lk} is only integration of Y_{lk} with signed measure $\rho r^l \mathrm{d}\lambda^3$ and the set $\{-k, -k+1, \ldots, 0, \ldots, k-1, k\}$ has for natural k only finite number of elements) we obtain desired equation

$$Q_{lm}(\Omega + \Delta\Omega) = \sum_{m'=-l}^{l} D^l_{mm'}(\Delta\Omega)^* Q_{lm'}(\Omega)\,. \tag{3.44}$$

Another beneficial transformation of multipole expansion is the ability to shift the origin of expansion. The following theorem (which shows, how to move the expansion into the origin of the coordinate system) is taken from [99] (theorem (5.3), adapted for our definitions) without proof.

Theorem 1. *Suppose that the distribution of charges is located inside the sphere D of radius a with center at $Q = (\rho, \alpha, \beta)$, and that for points $P = (r, \theta, \varphi)$ outside D, the potential due to these charges is given by the multipole expansion*

$$V(P) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{O_{nm}}{(r')^{n+1}} \sqrt{\frac{4\pi}{2n+1}} Y_{nm}(\Theta', \Phi')\,, \tag{3.45}$$

where $P - Q = (r', \Theta', \Phi')$. Then for any point $P = (r, \Theta, \Phi)$ outside the sphere D_1 of radius $a + \rho$

$$V(P) = \sum_{j=0}^{\infty} \sum_{k=-j}^{j} \frac{M_{jk}}{r^{j+1}} \sqrt{\frac{4\pi}{2j+1}} Y_{jk}(\Theta, \Phi)\,, \tag{3.46}$$

where

$$M_{jk} = \sum_{n=0}^{j} \sum_{m=-n}^{n} \frac{O_{j-n,k-m} i^{|k|-|m|-|k-m|} A^m_n A^{k-m}_{j-n} \rho^n \sqrt{\frac{4\pi}{2n+1}} Y_{n,-m}(\theta, \varphi)}{A^k_j}\,, \tag{3.47}$$

and where

$$A^m_n = \frac{(-1)^n}{\sqrt{(n-m)!(n+m)!}}\,. \tag{3.48}$$

By incorporating the rotation transformation into the equation (3.41) we get to [96, 97]

$$\begin{aligned} V = \sum_{l_1,l_2=0}^{\infty} \sum_{m'_1,m_1=-l_1}^{l_1} \sum_{m'_2,m_2=-l_2}^{l_2} \sum_{m=-l_1-l_2}^{l_1+l_2} [l_1, l_2] R^{-l_1-l_2-1} Q^{(1)}_{l_1,m'_1} Q^{(2)}_{l_2,m'_2} \\ (-1)^{l_1} D^{l_1}_{m_1,m'_1}(\Delta\Omega_1)^* D^{l_2}_{m_2,m'_2}(\Delta\Omega_2)^* Y_{l_1+l_2,m}(\Omega_{12}) \begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & m \end{pmatrix}, \end{aligned} \tag{3.49}$$

where the summation over m_1' and m_2' performs rotation of multipole expansions, summation over m_1, m_2 and m calculated energy contribution belonging to specific values of tuple (l_1, l_2). These contributions are then summed.

We would like to note a few things about the equations (3.39) and (3.49):

1. The formula is symmetrical with respect to swapping indexes 1 and 2. This follows from equation (3.17).

2. The calculated potential energy is real. This was one of the conditions during derivation.

3. It follows from equation (3.39) that multipole of order l can be encoded into $2l + 1$ independent complex numbers. However, since the relation (3.14) holds for all $l \in \mathbb{N}$ and $k \in \{-l, -l+1, \ldots, 0 \ldots, l-1, l\}$ we can use $2l + 1$ real numbers. Therefore, we define Q_{lkc} and Q_{lks} (c and s are just labels, not indices) by equations

$$\begin{aligned} Q_{lkc} &= \sqrt{\frac{1}{2}} \left((-1)^k Q_{l,k} + Q_{l,-k} \right) , \\ iQ_{lks} &= \sqrt{\frac{1}{2}} \left((-1)^k Q_{l,k} - Q_{l,-k} \right) , \end{aligned} \tag{3.50}$$

   where $k \in \{1, \ldots, l\}$ and for $k = 0$ this definition is basically useless, and we simply use Q_{l0} [6].

4. It might seem that we have to sum over 7 different indexes (and in equation (3.41) over 5 indexes). However, the delta function in formula (3.17) means that the summation actually runs only over 6 different indexes (and in equation (3.41) over 4 indexes)

5. When computing multipole energy up to order n, we sum over indexes l_1 and l_2 with condition $l_1 + l_2 \leq n$.

Our default strategy was to obtain coefficients Q_{lk} from integration of density matrix performed by existing DMA codes. To perform the expansion into distributed multipoles, one has to first select sites. Unfortunately, to our knowledge, there is no algorithm that would give us (in some sense) the best positions. The goal in selecting DMA sites is to minimize the displacement of the charge element. Typically, in ab-initio calculations, basis functions are centered around nuclei. Considering that nuclei also bear a charge that must be included in the intermolecular energy calculation, it is reasonable to view nuclei as natural expansion sites [26]. Overlap densities will be distributed along lines connecting

[6] Any Cartesian vector $\hat{V} = (V_1, V_2, V_3)$ can be identified with a vector in a spherical tensor formalism by

$$\hat{V}_0 = V_3 \, , \tag{3.51}$$

$$\hat{V}_{\pm 1} = \mp \frac{1}{\sqrt{2}} \left(V_1 \pm i V_2 \right) . \tag{3.52}$$

Therefore, the coefficients Q_{10}, Q_{11c}, and Q_{11s} are Cartesian dipole momenta. The functions Y_{lkc}, Y_{lks}, and Y_{l0} obtained in the same manner from spherical harmonics form a real basis in the space of real functions on $\mathbb{S}^2$ with the same orthogonality properties as spherical harmonics.

the nuclei of bonded atoms – sites positioned at the center of the bonds are a natural way to represent them [26]. By following this approach, no overlap charge element will need to be relocated more than a quarter of the bond length [26]. We will use this on the CO_2 molecule when trying to achieve good description of interaction by lower order multipoles.

In the case, when sites correspond to atoms, chemical literature gives us an interpretation of DMA [1]. Charges usually describe electronegativity effects. Dipoles describe atomic distortions and often arise from overlap of s and p orbitals. Quadrupoles can be associated with π bonds and are the sign of overlap of p orbitals. Higher-order terms are added for higher precision.

The algorithm integrating the density matrix has to decide, which element of charge will be added to which site. The used codes perform this by assigning a variable $\mathcal{R}_i$ to i-th site and selecting the site satisfying the condition

$$\min_i \frac{|\mathbf{r}_i|}{\mathcal{R}_i}\,, \tag{3.53}$$

where $|\mathbf{r}_i|$ is the distance of the discussed charge element from i-th site. When the sites correspond to atoms, the $\mathcal{R}_i$ is set by default to be 0.65 Å for all atoms except for hydrogen, where 0.325 Å is used. The utilized DMA codes use Lebedev angular quadrature, which is briefly introduced in the following section.

3.5 Lebedev angular quadrature

Most codes which aim to obtain chemical properties from ab-inito quantum chemical calculations have to evaluate an integral of the form

$$\int_{\mathbb{S}^2} f(\mathbf{x})\,\mathrm{d}\Omega = \int_0^{2\pi}\int_0^{\pi} f(\theta,\varphi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi\,, \tag{3.54}$$

where f is defined on $\mathbb{S}^2$.

Approaches using so-called Gauss quadratures approximate this integral by

$$\int_0^{2\pi}\int_0^{\pi} f(\theta,\varphi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \approx \sum_i w_i f(\mathbf{x}_i) \equiv Q[f]\,, \tag{3.55}$$

where w_i are some weights that have to be determined together with points $\mathbf{x}_i$. Since the theory behind one dimensional quadrature is well elaborated, one might use two-dimensional parametrization of the sphere and write

$$Q[f] = \sum_i \sum_j w_{ij} f(\theta_i,\varphi_j)\,, \tag{3.56}$$

where w_{ij} is usually evaluated with the help of the Jacobian of the function $(\theta,\varphi) \mapsto \mathbb{S}^2$. However, this does not have to give satisfactory results. Instead of applying products of one-dimensional Gauss quadratures, one may instead employ spherical harmonics to write

$$f(\theta,\varphi) \approx \sum_{l=0}^{l_{\max}} \sum_{m=-1}^{l} c_{lm}\, Y_{lm}(\theta,\varphi)\,, \tag{3.57}$$

(a) (b)

Figure 3.1: (a) Products of one-dimensional quadratures arising from $(\theta, \varphi) \mapsto \mathbb{S}^2$ with 225 points, and (b) Lebedev quadrature with 170 points. It is usually thought, that one needs $\frac{2}{3}$ of points from product quadrature to obtain the same precision using Lebedev quadrature. On the other hand, it is challenging to use symmetry properties with Lebedev points. Note that even when using Lebedev quadrature, all points don't have equal weights, which is a property of Chebyshev quadratures. The images were generated with the help of the numgrid [100, 101] Python package.

where coefficients c_{lm} can be in theory calculated by

$$c_{lm} = \int_0^{2\pi} \int_0^{\pi} f(\theta, \varphi)\, Y_{lm}^*(\theta, \varphi) \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\varphi \,. \tag{3.58}$$

A better option is to solve the system of linear equations (3.57) in some points on the sphere to find these coefficients. Since the integrals of spherical harmonics over the unit sphere are known (see equation (3.21)), the (3.54) can be simply found. We use a configuration of DMA with so-called Lebedev quadrature, which is implemented only for specific numbers of points on the unit sphere ($\{\ldots, 170, 194, 230, 266, 302, 350, 434, 590, 770, 974, 1202, \ldots\}$)

One, in general, needs the coefficients c_{lm} to vanish quickly with l to obtain an accurate value of integral (3.54) only by using l smaller than predefined threshold – the mathematical theory states, that the smoother the function f actually is, the better the approximation (3.57) with finite l becomes. In the DMA code, it is enough to obtain coefficients up to small order, since the physical nature of electrostatic force damps the higher orders of expansion (however, this doesn't necessarily mean that $l_{\max}$ in equation (3.57) can be small).

3.6 Optimization strategies

An alternative way to integration of a density matrix is to obtain coefficients Q_{lk} by fitting then on a precalculated dataset. This is done by making the difference (or other measure of similarity) between data and model relying on the coefficients Q_{lk} as small as possible by the means of optimization.

The unconstrained mathematical optimization problem is formulated by an equation (for simplicity, written for only one parameter)

$$Q(w) = \sum_{i=1}^{n} Q_i(w)\,, \tag{3.59}$$

where we wish to find the value of the parameter w which minimizes the cost/loss/objective function Q. The Q_i is many times selected as $|y_i - f(x_i)|^2$ which is obtained from the maximum-likelihood estimation for independent observations. For some specific functions f, there exist exact solution, that can be written in the closed form. However, even fitting dipole-dipole interaction on dimers is a quadratic problem with non-trivial constraint (when the complex spherical tensor for $l = 1$ is the unknown parameter, the constraints are given by requirement, that Q_{l0} and Q_{lkc} and Q_{lks} are real when calculated by equation (3.50)).

Therefore, the minimization is done by an iterative procedure (which we will call training in this text)

$$w^{(j)} = w^{(j-1)} - \eta \sum_{i=1}^{n} \nabla Q_i(w^{(j-1)}) \tag{3.60}$$

where the gradient can be calculated numerically, or sometimes explicitly in the closed form. The η parameter is called learning rate. The process described by equation (3.60) is called Gradient Descend and is implemented in many libraries (we have tried lsqfit package that can use scipy and GSL as a backend).

In our case, we had to work with noisy data; therefore, we had to use more elaborate methods. We have taken advantage of the TensorFlow library, which is used to train huge neural networks. There are three main advantages of this approach:

1. All operations performed by TensorFlow functions automatically calculate gradients, which means that we do not rely on possibly approximate evaluation from noisy calculations (we expect that this is the greatest obstacle for GSL library).

2. TensorFlow implements many optimization algorithms, that perform gradient descend. In our work, we use the Adam optimizer [102].

3. The structure of TensorFlow operations allows us to straightforwardly incorporate the constraint $Q_{l0}\,,\ Q_{lks}\,,\ Q_{lkc} \in \mathbb{R}$. The strategy will be described in section (4.3).

Adam calculates the moving average of the gradient and squared gradient, which are exponentially damped (this is controlled by parameters β_1 and β_2) [102]. The learning rate is updated based on these values separately for every parameter. Adam optimizer is part of the family of Stochastic Gradient Descend algorithms. This means that Adam performs step (3.60) calculated not from all data, but from a smaller sample, which should be beneficial in the case of noisy datasets. Therefore, the whole training procedure is performed as follows: Data are shuffled randomly and separated into batches[7]. The values $Q_i(w)$ from equation (3.59) are calculated for the whole batch (together with first derivatives and all needed gradients). Parameters are updated by the equation (3.59). Subsequently, learning rates for all parameters are updated. The process of evaluating all batches is called an epoch, which is repeated many times.

[7]Evaluation by batches should be faster and less prone to errors in separate datapoints and noise.

The whole training procedure relies on several values (also called hyperparameters) which are, in our case, only parameters of the optimizer and initial guesses of the multipole coefficients. To find those, one can easily utilize GridSearch, which is simple looping over all combinations and selecting the one with the best results.

4. Results

We have developed a code (see appendix (B)) that can obtain coefficients Q_{lkc}, Q_{lks}, and Q_{l0} from the integration of ab-initio density matrix performed by the Molpro DMA or GDMA programs. We have implemented equation (3.41) and the rotation of coefficients from the equation (3.44). Furthermore, we have tested the stability of the multipole expansion with respect to the basis set and utilized ab-initio method. Moreover, we have investigated the quality of description of electrostatic interaction in rather close regions and the consistency of rotational transformation of multipole expansion with multipole expansion calculated from the rotated density matrix. These results are in section (4.1). Comparisons of electrostatic binding energies calculated from DMA with reference SCF calculations were done for structures with different physical and chemical properties in section (4.2). In section (4.3), tests were performed, whether it is possible to fit coefficients of multipole expansion (single multipole on one molecule) on dimer benchmark data. To verify whether fitting can be used for good description of distributed properties, we have tried to fit the simplest model of correlation energy contributions using the tools required for the fitting of the coefficients of the multipole expansion.

We have used Molpro [80] and Psi4 [103] software packages for ab-initio calculations. Both of these software products work as a collection of tools and runtime environments where other authors implement their quantum chemistry codes. Psi4 is claimed to be faster than Molpro since it can utilize parallel environments better. Molpro, on the other hand, emphasizes accuracy and precise treatment of correlation effects, utilization of localized orbitals, explicitly correlated and multireference methods. The advantage of the Psi4 package is that it is effortless to use. It can be installed in an anaconda environment and called directly from python scripts. However, the biggest motivation to use this community-driven package is its ability to produce `fchk` files with density matrix (format commonly used by Gaussian ab-initio codes). In the table (4.1), we have summarized configurations of our calculations that were performed by those two codes. The coefficients of multipole expansion were calculated using Molpro DMA code (which gave us rather unsatisfactory values) and GDMA[1] using `fchk` files generated by Psi4 ab-initio calculations.

In the following section, configurations A, B, C, and D from table (4.1) were used for the calculation of density matrices, which were expanded into the multipole coefficients. The configurations A, C, and E were used for the ab-initio reference binding curves. We have used the default configuration of Molpro DMA code (it was challenging to find documentation for this code and apart from the specification of density matrix and order of the expansion, the options did not seem to work). In GDMA, the results are for the default numerical settings – using Lebedev angular quadrature with 590 points. We tried changing this configuration for structures, where GDMA performs poorly. The GDMA configuration is in appendix (B) in listing (B.4). Finally, the configuration F from table (4.1) was applied for the calculation of the dataset for fitting the coefficients Q_{lk}.

[1] To prevent future confusion, GDMA and Molpro DMA are the codes for the integration of the density matrix, whereas DMA is the multipole expansion.

Table 4.1: Computational details – only configurations explicitly altered from their default settings are listed. `DF_BASIS` and `DF_BASIS_EXCH` is the basis used for density fitting (described in section (2.5)). The parameter `scale_trip` gives a coefficient for obtaining explicitly correlated triple contributions, as described in section (2.6).

Configuration	Package	Description
A	Molpro	SCF energy=1.d-16,twoint=1.d-19,oneint=1.d-19, zero=1.d-19,orbital=1.e-11,thrao=1.d-19, thrmo=1.d-19,thrsw=1.d-19,throv=1.d-19, thrprod=1.d-19, accu=1.d-17
B	Psi4	SCF with default configuration
C	Molpro	CCSD(T) the same configuration as A
D	Psi4	CCSD(T) with default configuration
E	Molpro	SAPT0 with the same configuration as A
F	Molpro	CCSD(T)–F12B in AVTZ basis DF_BASIS=AVQZ, DF_BASIS_EXCH=AVQZ/JKFIT scale_trip=1

4.1 Stability and convergence of multipole expansion

In this section, we present results regarding stability of multipole expansion with respect to the cardinal number of basis set, as well as how the multipole expansion converges towards the ab-initio binding curve with growing order of expansion.

We have started by plotting binding curves for methanol, hydrogen fluoride, and CO_2 with coefficients calculated by Molpro DMA code (some examples are in figures (4.1) and (4.3)); however, we have not obtained satisfactory results. The binding curves calculated from multipole expansion did not correspond well to the results from ab-initio calculations. For methanol, we have obtained good description only in the third order for the aug-cc-pVQZ basis (see figure (4.1)). Different basis and different orders, on the other hand, gave completely meaningless results (especially aug-cc-pVDZ and cc-pV5Z). The coefficients obtained from Molpro DMA code completely fail to describe binding curves of hydrogen fluoride (see figure (4.3)), where we might see something like convergence regarding the size of augmented Dunning correlation consistent set. However, even in the biggest basis set available, the description is not satisfactory. For the CO_2 molecule, we have not been able to get any meaningful results. We observed that all results are highly dependent on the basis of the density matrix, which was not expected at all. Especially the behavior of methanol, where the description with multipole coefficients obtained from (aug)-cc-pV5Z basis was as bad as from (aug)-cc–pVDZ basis, is puzzling. One might argue that we are in a region, where SCF calculation includes different effects than only electrostatic contributions. However, we have also calculated the CCSD(T) binding curve, and we have checked that the electrostatic term in SAPT0[2] calculation is much closer to the SCF energy than DMA with coefficients obtained from Molpro DMA code, thus ensuring that the comparison of DMA with SCF is meaningful.

After these findings, we have changed the code that was used for the integration of density matrix (i.e., obtaining coefficients of expansion) and started using GDMA (we have also had to change ab-initio software to Psi4, since Molpro cannot produce `fchk` files). We have plotted several coefficients of multipole expansion depending on basis, molecule, DMA code, and used ab-initio method (see figures (4.4), (4.5), (4.6) and (C.2)). We may notice, that even the zeroth order of expansion (charge) is unstable when using Molpro DMA code.

Contrary to the Molpro DMA code, we have obtained stable results in the zeroth and first order of expansion by using the GDMA. When comparing stability of coefficients from Molpro DMA code and from GDMA, we might for some structures and some coefficients observe slow convergence of Molpro DMA solution towards the GDMA coefficients with respect to the basis set size (see figure

[2]Symmetry adapted perturbation theory utilizes an alternative strategy to supermolecular approach obtained from equation (2.58). The SAPT0 computes intermolecular energy from SCF densities by perturbation theory by utilizing an equation

$$E_{\text{int}}^{\text{SAPT0}} = E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)} + E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)}\,, \tag{4.1}$$

where we are only interested in the electrostatic term $E_{\text{elst}}^{(1)}$. This approach does not suffer from BSSE and by using higher orders of expansion it might compete with CCSD accuracy.

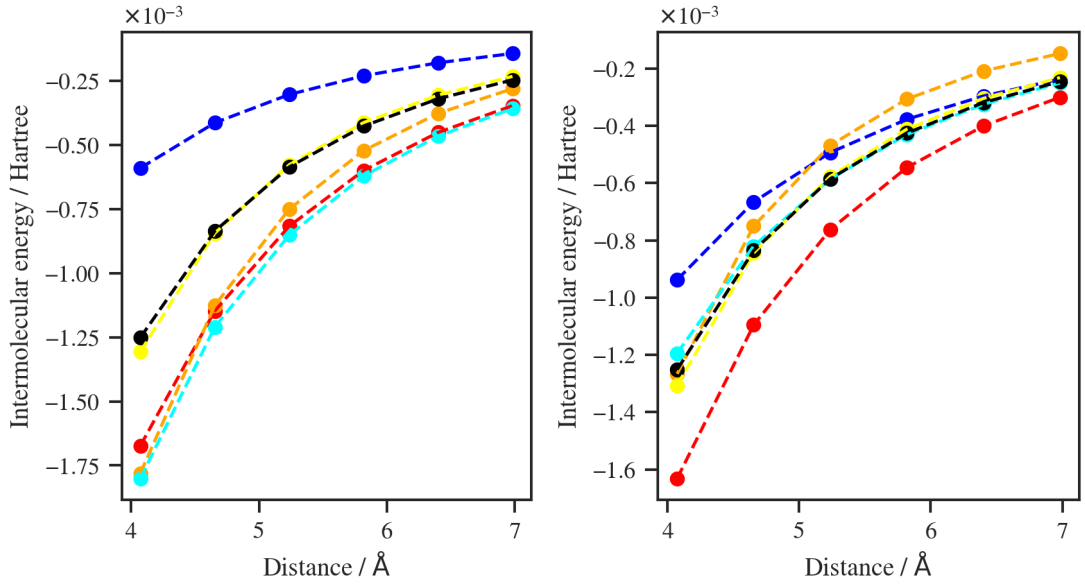
(4.5)). This is not the case for methanol (see figure (4.4)) on the other hand. We have also tested higher-order terms, where even GDMA does not yield identical results for all basis sets; however, we may see that the coefficients might converge in bigger basis (see figures (4.7) and (C.3) – (C.9)). This is in contrast with Molpro code, where there is no convergence at all. It is interesting that coefficients obtained from aug-cc-pVnZ seem to converge from one side and coefficients from cc-pVnZ basis from the other (see the "zigzag" pattern in figures (4.7) and (C.3) – (C.9)). The binding curves from DMA are good approximation to those obtained from ab-initio calculations, even for relatively close molecules for methanol and hydrogen fluoride. Furthermore, adding higher terms into expansion did produce better results and the intermolecular energy did converge well towards the SCF reference (see figure (4.2) or (C.1) for example).

The exception to this rule is the CO_2 molecule, where coefficients obtained from GDMA gave us completely useless short range binding curves (up to third order of multipole expansion). The fourth order has usually improved the results significantly, as we will demonstrate in the next section. It is possible to see in figure (4.6) that the coefficients from Molpro DMA code and GDMA are entirely different even in the zeroth order of expansion. Changing the parameters of Lebedev quadrature within the GDMA code did not produce significantly better results. Other strategies will be presented in the next section.

To finish the study of stability of multipole expansion coefficients, we have generated 30 randomly rotated methanol molecules in the aug-cc-pV5Z basis for both Molpro DMA and GDMA codes and compared them with rotated expansion of non-rotated molecule. The most significant difference is at the fourth digit, and it is challenging to claim which numerical effect is responsible for it. The results are in figure (C.10)).

To verify that our implementation is valid and all problems with definitions of special functions as well as Euler angles and quaternions were successfully resolved, we have made several elementary tests. Foremost, we have checked that even in the situation where there is no truncation of imaginary parts of complex values, we obtain from the equation (3.41) real energy and that when the molecules 1 and 2 in the same equation are flipped, we obtain the same value. We have also checked, that we obtain the same value for all symmetrical terms in the benchmark for methanol dimers when using the GDMA coefficients.

From the presented plots of coefficients of multipole expansion (figures (4.4), (4.5), (4.6), and (C.2)) it appears that they do not depend much on the ab-initio method that was used to obtain the density matrix. This dependence will be investigated in the next section by comparing binding energies calculated from DMA with benchmark calculations.



(a) aug-cc-pVTZ (b) aug-cc-pVQZ

Figure 4.1: Binding curves for the hydrogen bonded methanol dimers where the x axis shows distance from equilibrium position in the methanol crystal– DMA of different orders with coefficients obtained from SCF calculation in Molpro on aug-cc-pVTZ/aug-cc-pVQZ basis, compared with SCF and CCSD(T) calculations from Molpro (BSIE correction from aug-cc-pVTZ and aug-cc-pVQZ basis). For legend, see Figure (4.2).

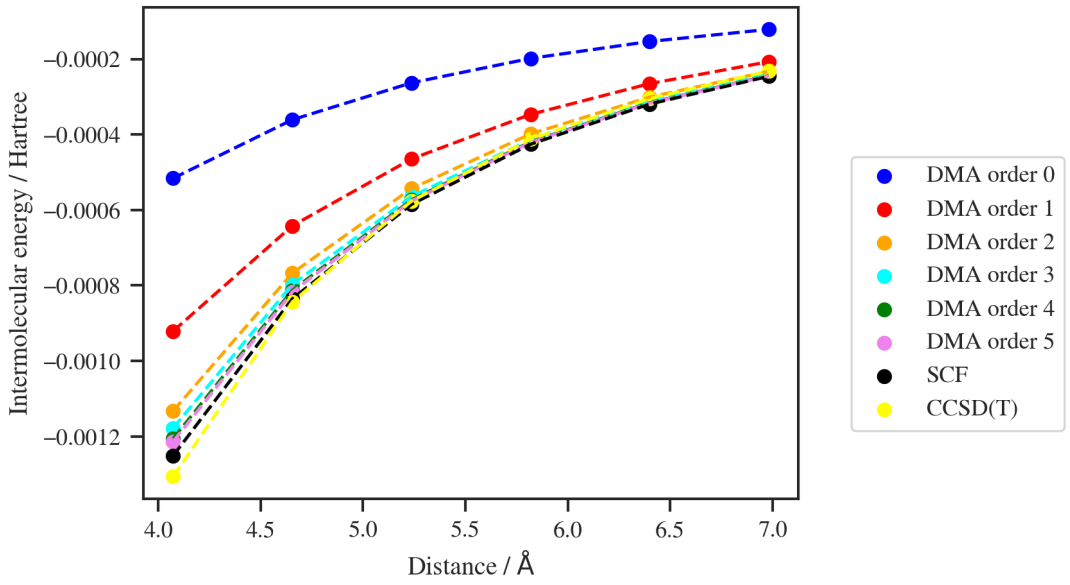
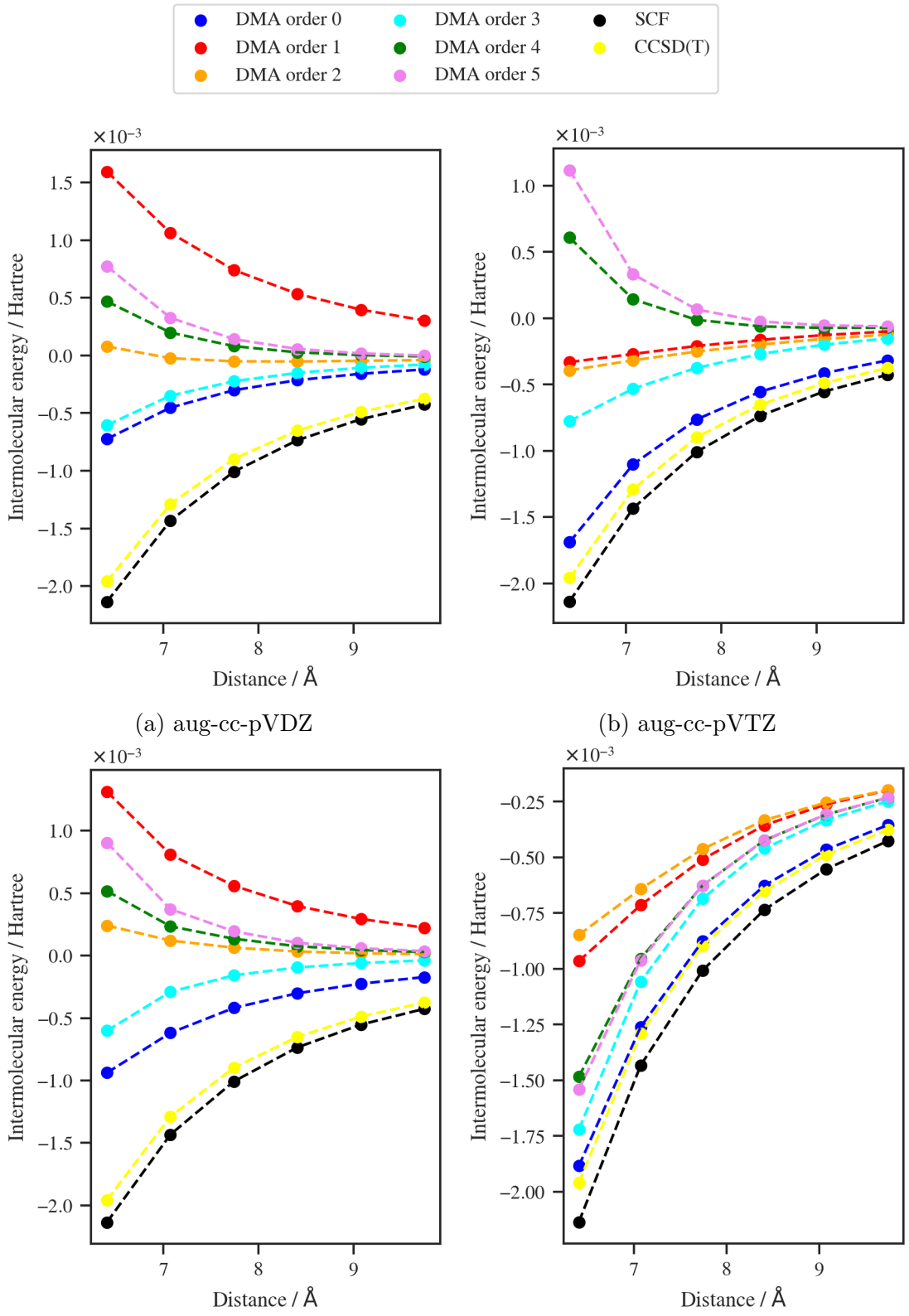


Figure 4.2: Binding curves for the hydrogen bonded methanol dimers where the x axis shows distance from equilibrium position in the methanol crystal – DMA of different orders with coefficients obtained from SCF calculation in Psi4 and GDMA on aug-cc-pV5Z basis, compared with SCF and CCSD(T) calculations from Molpro (BSIE correction from aug-cc-pVTZ and aug-cc-pVQZ basis).



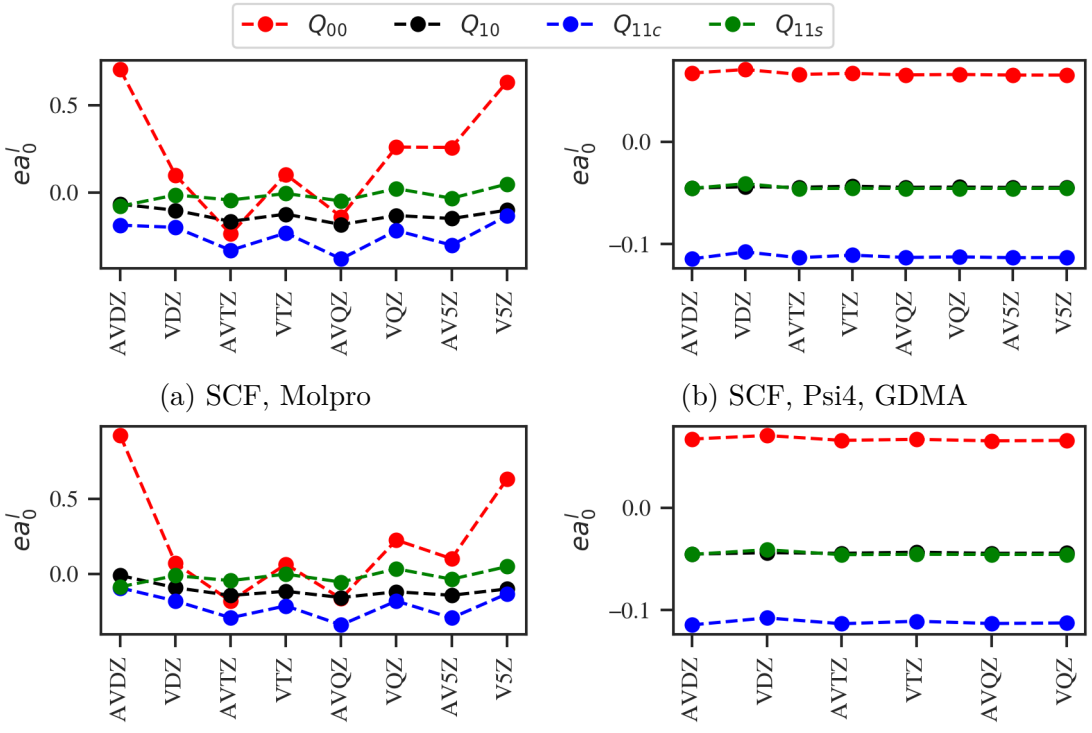
(a) aug-cc-pVDZ

(b) aug-cc-pVTZ

(c) aug-cc-pVQZ

(d) aug-cc-pV5Z

Figure 4.3: Hydrogen fluoride binding curves – DMA of different orders with coefficients obtained from SCF calculation in Molpro on aug-cc-pVDZ/aug-cc-pVTZ/aug-cc-pVQZ/aug-cc-pV5Z basis, compared with SCF and CCSD(T) calculations from Molpro (BSIE correction from aug-cc-pVQZ and aug-cc-pV5Z basis). The distance between F atoms is shown.



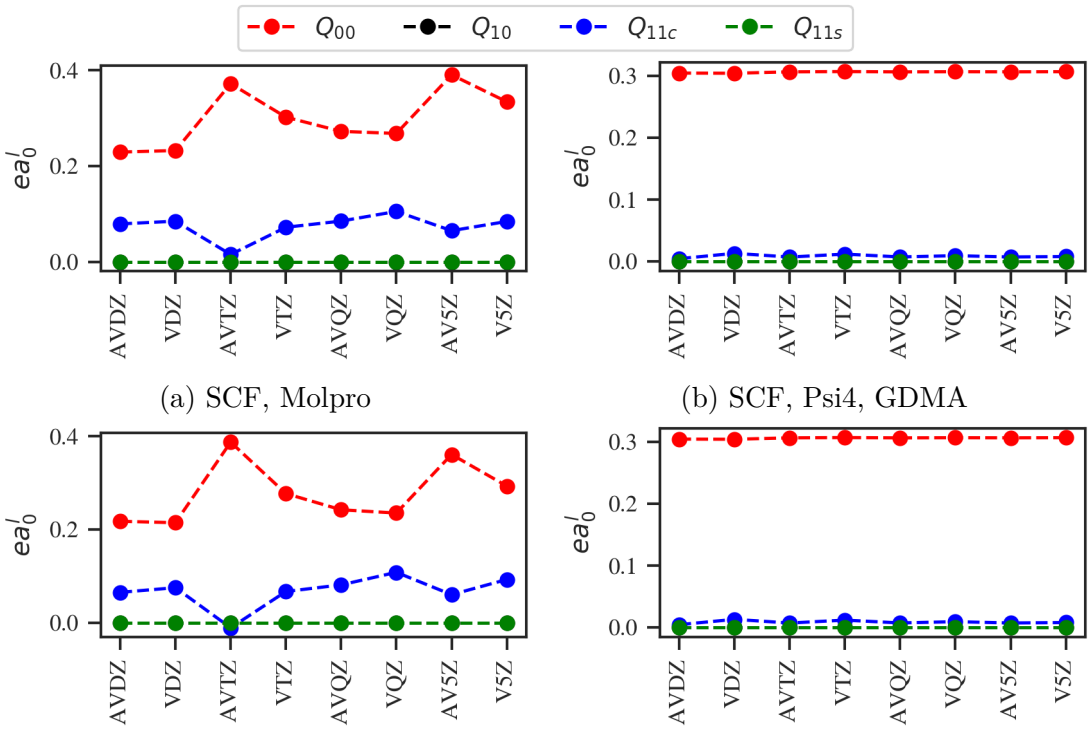
(a) SCF, Molpro

(b) SCF, Psi4, GDMA

(c) CCSD(T), Molpro

(d) CCSD(T), Psi4, gdma

Figure 4.4: Methanol: effects of basis, DMA code, and ab-initio method in the zeroth and first order of expansion. The y-axis units depend on coefficient order.



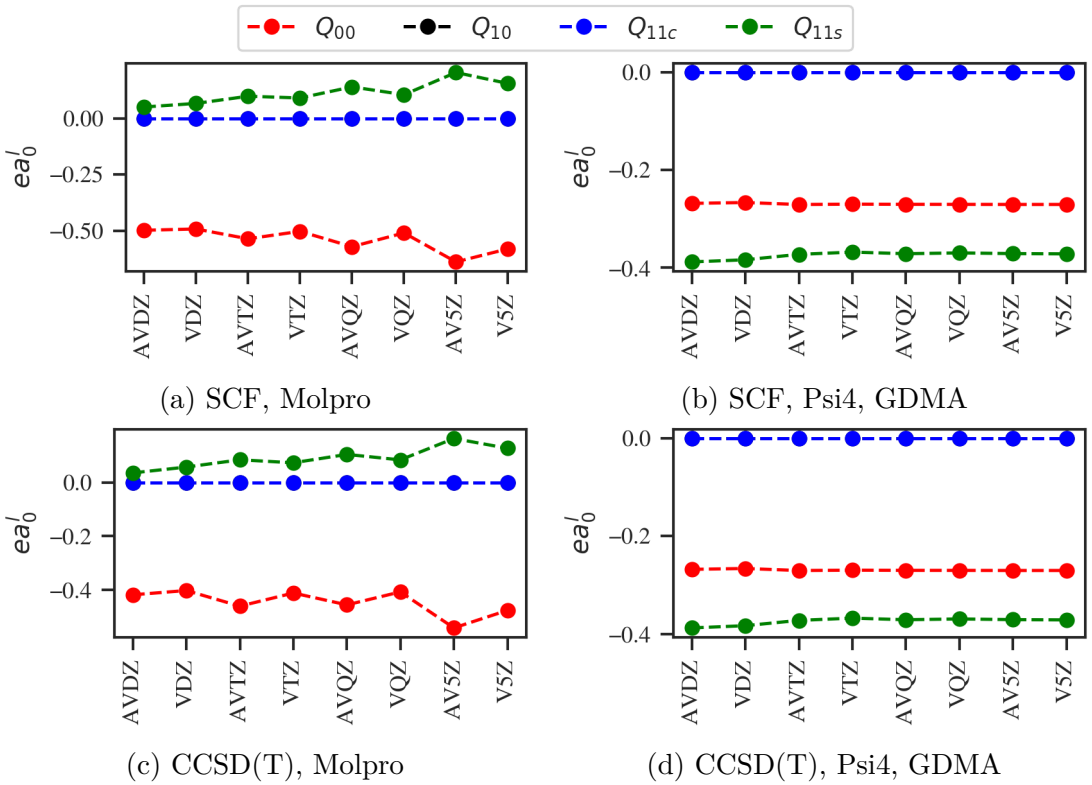
(a) SCF, Molpro

(b) SCF, Psi4, GDMA

(c) CCSD(T), Molpro

(d) CCSD(T), Psi4, GDMA

Figure 4.5: Hydrogen fluoride: effects of basis, DMA code, and ab-initio method in the zeroth and first order of expansion. The y-axis units depend on coefficient order.



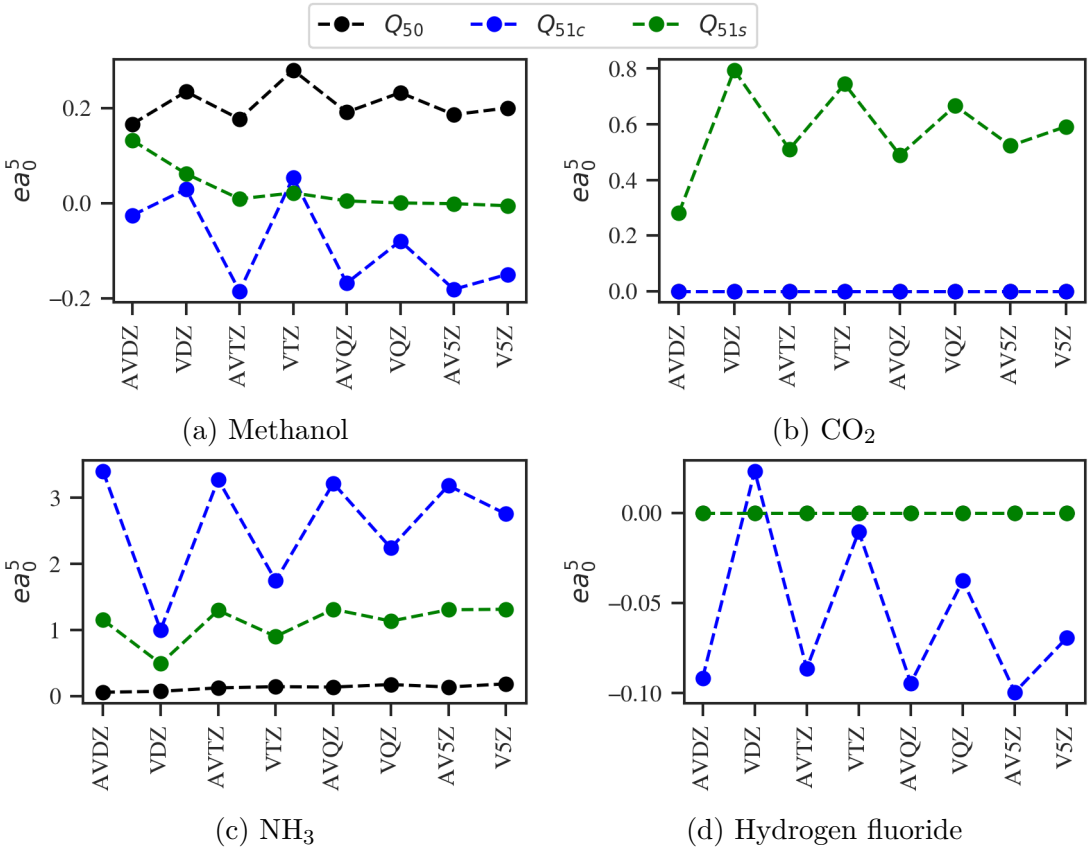
(a) SCF, Molpro

(b) SCF, Psi4, GDMA

(c) CCSD(T), Molpro

(d) CCSD(T), Psi4, GDMA

Figure 4.6: CO_2: effects of basis, DMA code, and ab-initio method in the zeroth and first order of expansion. The y-axis units depend on coefficient order.



(a) Methanol

(b) CO_2

(c) NH_3

(d) Hydrogen fluoride

Figure 4.7: Multipoles at 5th order from SCF density matrix in Psi4 and GDMA.

4.2 Testing electrostatic contributions on molecular solids

In this section, we compare our DMA implementation with calculations from paper [17] where the binding energies of methanol, methane, CO_2, and HN_3 were obtained mostly by SCF and MP2 calculations (in the fragmented approach). The fragmented approach was performed by considering all dimers in a box around one central molecule, meaning, that there are only a few dimers with high intermolecular distance and most dimers have distance somewhere in the middle (this will be relevant for the discussion of methane, where most common distance is between 20 Å and 30 Å). The dataset for this paper is publicly available at [104]. Since we are dealing with molecular solids, the most natural variable to compare is the contribution of dimers to the binding energy of the crystal (in our case backward calculated cumulative sum of selected SCF/DMA dimer contributions).

We present results in two forms. The first set of figures show backward calculated cumulative sums of DMA (up to third order with coefficients obtained from GDMA using SCF density matrix), SCF, as well as MP2. As a second way to evaluate DMA accuracy, we plotted the absolute value of the difference between backward calculated cumulative sum of SCF and DMA in different orders of expansion. The regions for the second set of plots were carefully selected to allow meaningful comparison.

For the second set of plots, we have also compared results numerically in table (4.3) – for coefficients of multipole expansion obtained from SCF procedure – and table (4.4) – for coefficients of multipole expansion obtained from CCSD(T) calculations. Description of the datasets for plots showing the difference between SCF and DMA backward calculated cumulative sums and tables (4.3), (4.4) is presented in the table (4.2).

For methanol, we have truncated data in figure (4.8) to include only dimers with distance less than $60a_0$ – this means that the comparison included 2667 dimers. In the figure (4.8), we see good correspondence between SCF and DMA. For the comparison of different orders, we observe in figure (4.9) that the description becomes better (on average) by introducing higher-order DMA terms, but in the sum over the selected region higher-order terms are not significantly better than the lower-order ones. This can also be seen in tables (4.3) and (4.4), where even the zeroth order of expansion seems to give reasonable results.

Table 4.2: Description of datasets for comparison in tables (4.3), (4.4) and plots (4.9), (4.11), (4.13), and (4.15). Technical information can be found in [17, 104]. The basis were selected according to numerical inaccuracies presented in distant dimer contributions after consulting [104].

Crystal	basis	closes / a_0	most distant / a_0	number of dimers
methanol	AVTZ	15.9	54.4	1946
CO_2	AVDZ–AVQZ fit	13.0	49.8	153
NH_3	AVTZ	25.1	95.4	1449
methane	AVQZ	20.6	28.1	19

Table 4.3: Comparison of the backward calculated cumulative sum obtained from DMA (using the coefficients from SCF method on V5Z basis) and from ab-initio calculations. Datasets in table (4.2) were used. Values are in Hartree.

	reference	DMA up to order				
Crystal	SCF	0	1	2	3	4
methanol	-0.00118	-0.00111	-0.00120	-0.00126	-0.00126	-0.00125
CO_2	-7.03E-5	-3.24E-5	1.12E-4	1.14E-4	-5.45E-5	-7.16E-5
NH_3	1.62E-5	1.49E-6	7.91E-6	1.47E-5	1.44E-5	1.34E-5
methane	1.57E-8	-4.19E-10	8.76E-9	-4.85E-8	2.30E-8	6.92E-8

Table 4.4: Comparison of the backward calculated cumulative sum obtained from DMA (using the coefficients from CCSD(T) method on V5Z basis) and from ab-initio calculations. Datasets in table (4.2) were used. Values are in Hartree.

	reference	DMA up to order				
Crystal	SCF	0	1	2	3	4
methanol	-0.00118	-0.00111	-0.00119	-0.00125	-0.00125	-0.00124
CO_2	-7.03E-5	-3.24E-5	1.12E-4	1.14E-4	-5.45E-5	-7.16E-5
NH_3	1.62E-5	1.50E-6	7.92E-6	1.47E-5	1.44E-5	1.34E-5
methane	1.57E-8	-4.19E-10	8.76E-9	-4.85E-8	2.30E-8	6.93E-8

Another satisfactory result was obtained for the NH_3 crystal. From figure (4.12), we see that the description of electrostatic binding energy is almost identical with the SCF calculation. The comparison of different orders in figure (4.13) shows us that the introduction of the first and the second order of expansion makes results significantly better when compared only with charges. From tables (4.3) and (4.4) we might conclude that the best description of intermolecular energy is at second order for the NH_3 molecule. For smaller orders of expansion, the results are of an order of magnitude smaller than reference calculations.

For methanol and ammonia, the total dimer energy is composed of a large number of positive and negative contributions. Even in this case, the precision of DMA seems satisfactory, and the SCF energies can be reasonably approximated by DMA with GDMA coefficients.

In figure (4.10) we may observe that DMA for CO_2 molecule is not precise enough by using only the third order of expansion; yet, it was expected when considering discussion from the previous section. Since contributions higher than $40a_0$ are negligible, the numerical comparison was made for dimers closer than this value. In figure (4.11), the fourth and fifth orders of DMA are shown to be essential. In the previous section, we have already tried to make low order coefficients of CO_2 molecule more precise by changing parameters of Lebedev quadrature. For the comparison on molecular solids, we have also tried to add virtual sites in the middle of chemical bonds. Interestingly enough, this did not make the description more precise and the results were actually worse (again up to the third order of expansion). Since the CO_2 molecule is known for its strong quadrupole moment, the importance of the fourth order of expansion (as an order that introduces quadrupole-quadrupole interaction) when using only

one multipole on molecule is expected, however it appears that the quadrupole cannot be, in this case, well described by lower order distributed expansion. This is problematic, since the main advantage of distributed treatment is the ability to use lower orders of expansion. The fourth order of expansion produced by GDMA code is not stable with respect to utilized basis set in ab-initio calculations.

The nonpolar methane molecule is rather exceptional. We may see an interesting pattern in figure (4.14) for the SCF data. The growth in electrostatic energy in figure (4.14) for distances higher than $30a_0$ is given by numerical inaccuracies that are present in the ab-initio calculations. The derivative of the SCF curve is given by the number of dimers that are in the considered distance. We expect that (with exception to really small distances) the DMA results are more reliable than ab-initio calculations for the electrostatic contributions. This has made the numerical comparison between SCF and different orders of DMA rather challenging. We have selected dimers with distance between $20a_0$ and $30a_0$, however, this is a rather arbitrary choice. It is shown in figure (4.14) that even for the stronger short range contributions, the result of third order of expansion is not a good approximation of SCF calculation. The figure (4.15) then shows rather weird behavior of different orders of expansion. However, it is a question whether bigger inaccuracies are in calculated coefficients or in SCF binding energies.

We may also observe, by comparison of tables (4.3) and (4.4), that the usage of CCSD(T) method instead of SCF has only a minor influence on the energies calculated by DMA. It is not clear why. In the previous section, one might find small differences in figures (4.4) – (4.7) between coefficients obtained from SCF and CCSD(T) calculations, ensuring that the command (B.1) in appendix (B) which should save CCSD(T) density matrix does not save SCF density matrix instead.

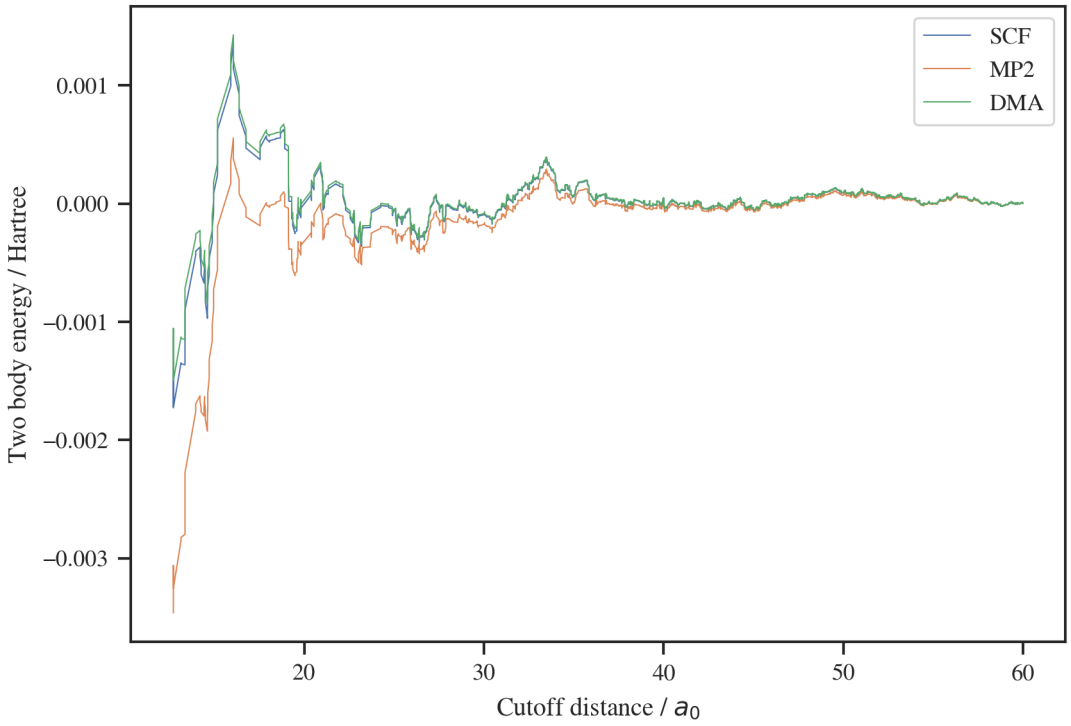


Figure 4.8: Backward calculated cumulative sum of SCF energy, total MP2 energy, and the energy obtained from DMA (GDMA up to third order from SCF density matrix) for methanol cystal.

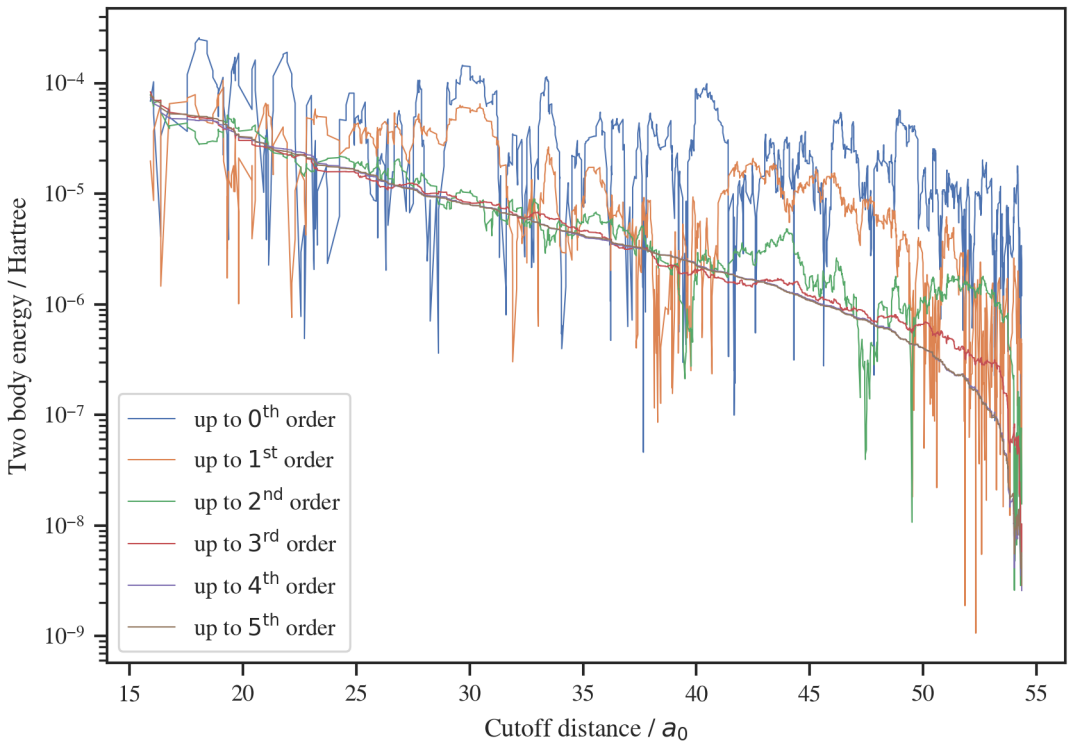


Figure 4.9: Comparison of absolute value of difference of backward calculated cumulative sum of DMA of different orders and backward calculated cumulative sum of SCF for methanol crystal.

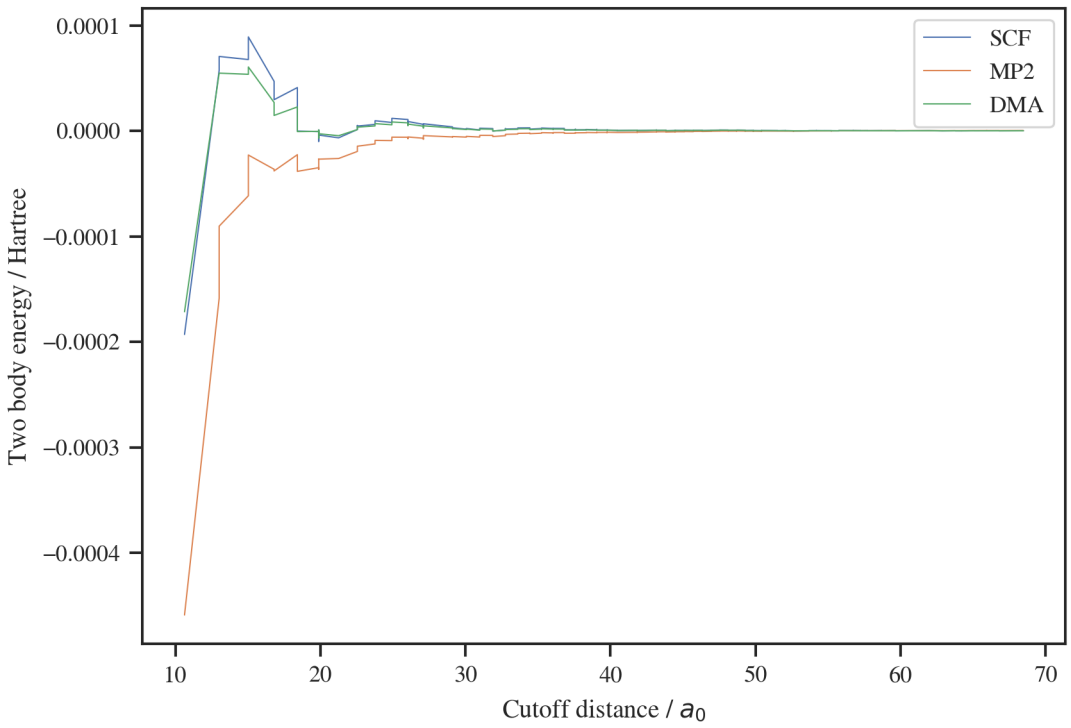


Figure 4.10: Backward calculated cumulative sum of SCF energy, total MP2 energy, and the energy obtained from DMA (GDMA up to third order from SCF density matrix) for CO_2 crystal.

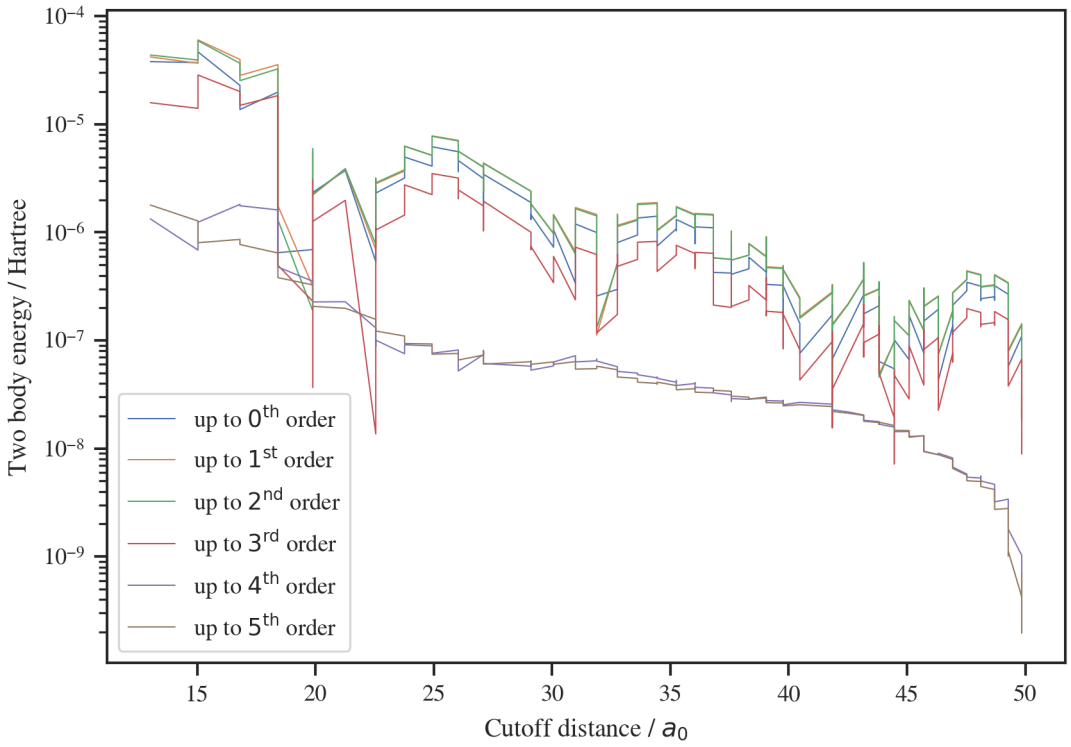


Figure 4.11: Comparison of absolute value of difference of backward calculated cumulative sum of DMA of different orders and backward calculated cumulative sum of SCF for CO_2 crystal.

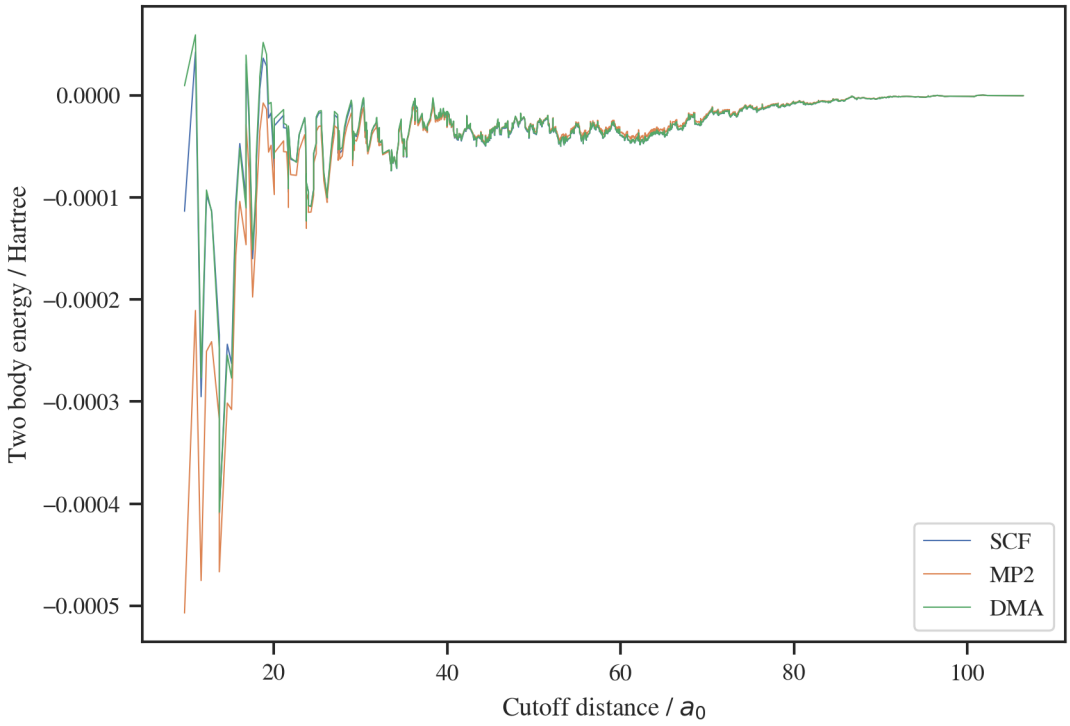


Figure 4.12: Backward calculated cumulative sum of SCF energy, total MP2 energy, and the energy obtained from DMA (GDMA up to third order from SCF density matrix) for NH_3 crystal.

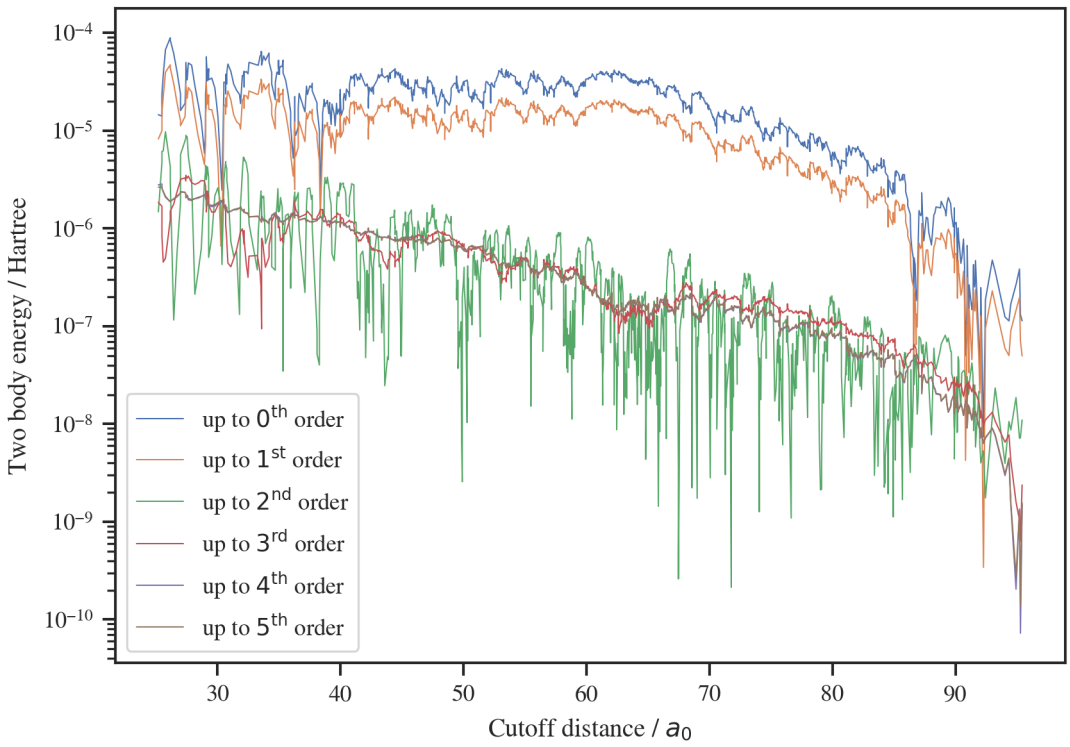


Figure 4.13: Comparison of absolute value of difference of backward calculated cumulative sum of DMA of different orders and backward calculated cumulative sum of SCF for NH_3 crystal.

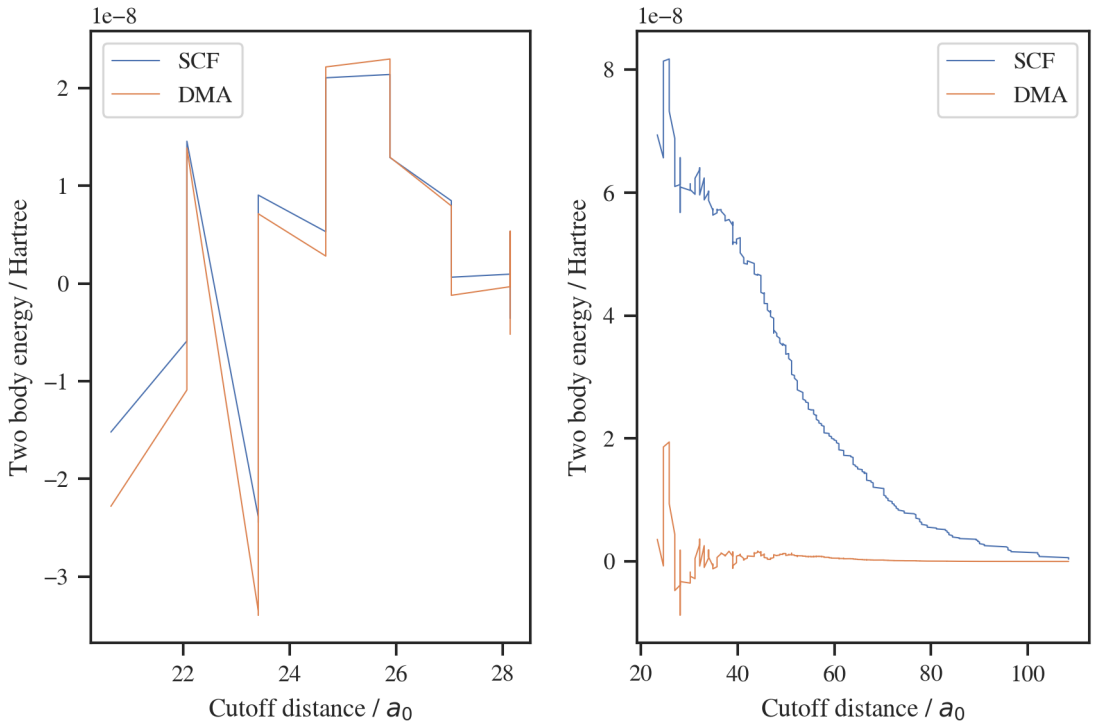


Figure 4.14: Backward calculated cumulative sum of SCF energy and the energy obtained from DMA (GDMA up to third order from SCF density matrix) for methane crystal. Since the MP2 energy is of order of magnitude larger, we do not plot it here. The left plot is backward calculated cumulative sum on the same region as plot (4.15).

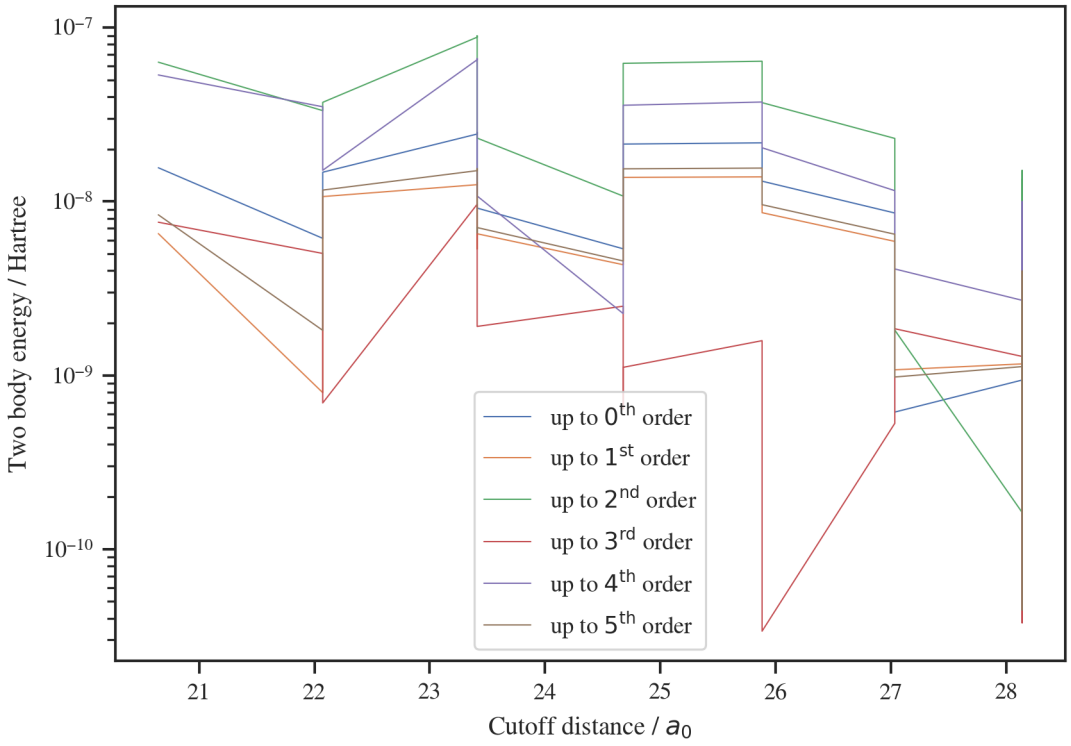


Figure 4.15: Comparison of absolute value of difference of backward calculated cumulative sum of DMA of different orders and backward calculated cumulative sum of SCF for methane crystal.

4.3 Fitting coefficients of multipole expansion

In this section, we use data calculated by the configuration F from Table (4.1) (explicitly correlated CCSD(T)-F12b calculations of methanol dimers in Molpro). The original goal was to fit models of correlation energy; however, our data contained many electrostatic contributions. Since the coefficients of multipole expansion from GDMA do not change much when CCSD(T) is used instead of SCF, we needed to get rid of the electrostatic energy by different means. The idea was to fit dipole-dipole interaction on distant dimers, remove these contributions from the data, and fit the correlation model on closer dimers. This ultimately did not work. However, we have still been able to obtain interesting results (even for the usage of DMA and possibilities of fitting distributed properties) which we will summarize here.

We have implemented a Keras model describing the dipole-dipole interaction. A slightly simplified version of this model is implemented in the listing (B.5). Let us describe the fitting procedure in the TensorFlow library. We first define real parameters Q_{10}, Q_{11c} and Q_{11s} (so-called weights in TensorFlow terminology) and by using the equations (3.50) we transform them into the complex vector Q. Subsequently, the two-dimensional tensor[3] $Q' = QQ^T$ is obtained. The precalculated values of spherical harmonics are the independent variables, since they are costly to evaluate, and it is really challenging to calculate them using only the operations allowed by TensorFlow. They are given to the model in the form of a two-dimensional tensor whose components are obtained as

$$\begin{aligned} I_{m_1,m_2} = \sum_{j=-l_1-l_2}^{l_1+l_2} & [l_1,l_2] R^{-l_1-l_2-1} (-1)^{l_1} \\ & Y_{l_1+l_2,j}(\Omega_{12}) \begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & j \end{pmatrix}, \end{aligned} \tag{4.2}$$

which is basically the equation (3.41) without the summation over m_1, m_2, l_1, and l_2. By elementwise multiplication of tensors I and Q' and summation over dimensions, we obtain the energy of dipole-dipole interaction for one data point.

Our implementation of the whole equation (3.49) using only the TensorFlow operations, which would be independent of the order of the expansion, was either flawed or the fit is much more challenging for the TensorFlow library than we have expected. It was built by generalization of the presented procedure for different l_1 and l_2 and summation of all the resulting terms with different (l_1, l_2) together (up to condition $l_1 + l_2 \leq n$ where n is the order of expansion).

Now, let us discuss the results of dimer-dimer fit. When employing the lsqfit package with scipy backend, we have not been able to perform the fit. The results from GSL library had averaged mean absolute percentage error around 80% whereas the TensorFlow model (which is in Figure (4.16)) performed much better with the same metric having a value around 16%. The only disadvantage was that the result depended on the initial values of coefficients a lot. Grid search is necessarily needed if higher-order terms were added to the model.

To test, how well a TensorFlow model can capture the distributed nature of multipole expansion, we have employed an entirely different task. We have used

[3] In this section, by tensor we mean matrix of several dimensions, which is terminology from TensorFlow.

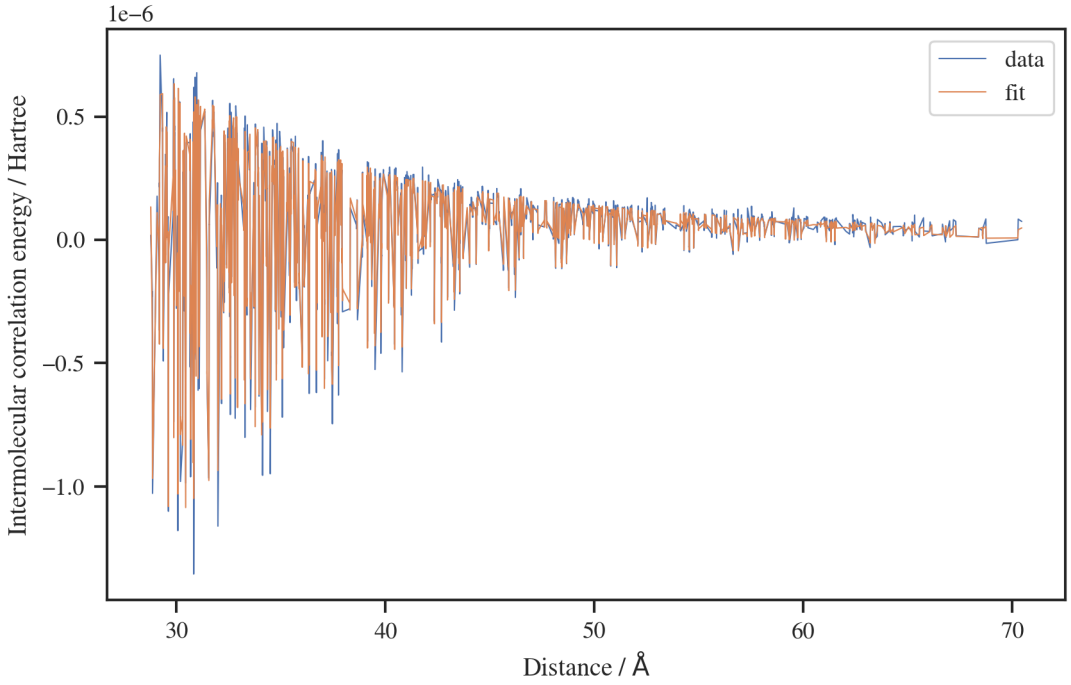


Figure 4.16: Fit of dipole-dipole interaction in long-distance correlation energy from CCSD(T)-F12b dimer data using the TensorFlow library and the dipole-dipole part of the Keras model (B.5). We have removed the CABS contributions from the correlation energy, since they were giving unreasonably large values.

the simplest model of correlation energy, which may be written as

$$E_c = \sum_{ij} \frac{C_{6ij}}{|\mathbf{R}_i^{(1)} - \mathbf{R}_j^{(2)}|^6}\,, \tag{4.3}$$

where C_{6ij} are the unknown coefficients and $\mathbf{R}_i^{(l)}$ is the position of i-th atom in l-th molecule. The advantages of this task are that this model is simple enough (in fact, it is linear in the parameters C_{6ij}) to be solvable by pseudoinverse; on the other hand, similarly to distributed multipoles, all the distances in the fraction in equation (4.3) are of almost the same value. By employing the lsqfit package, we have not obtained any meaningful result (this time not even with the GSL backend). In this case, even TensorFlow failed to obtain a solution close to the one obtained by the pseudoinverse calculation. One might, of course, argue that, according to the results of dimer-dimer fit in figure (4.16), the correlation contributions described by equation (4.3) is not much present in our dataset, however the best solution in the sense of least-squares exists. We have also tried to perform the fit of function (4.3) when we subtracted the dipole-dipole contribution from the dataset. This did not change the results at all.

Based on these attempts, we expect that obtaining coefficients of distributed multipole expansion from the dimer datasets is a really hard problem due to many existing local minima. On the other hand, fitting one multipole on a molecule with coefficients of multiple orders of expansion is probably achievable but rather hard from the coding perspective. The space of initial parameters has to be diligently scanned by GridSearch or other optimization algorithm.

Conclusion

We have studied two codes for obtaining coefficients of multipole expansion. We have shown that the method implemented in Molpro cannot be trusted since the results are highly dependent on the size of the basis set. On the other hand, the GDMA code gives results that are stable regarding the used basis for small (≤ 3) orders of expansion and for some structures even higher-order terms seem to be reliable. When comparing these two pieces of software, one cannot miss that unstable coefficients always appear to converge in bigger basis when calculated from the GDMA code (yet basis that is not available in ab-initio packages). That is not always the case for results from Molpro DMA code.

For methanol and hydrogen fluoride, we were able to describe binding curves accurately by distributed multipole expansion when compared with SCF calculations even by using low orders (up to three). A different situation is for the CO_2 molecule, where low orders with coefficients obtained from GDMA did not describe interaction well, however fourth and fifth order did improve results significantly. However, the CO_2 molecule is still problematic, since there are considerable differences between coefficients from Molpro DMA and GDMA code even in the zeroth order of expansion.

Our findings are not in complete agreement with the literature. We refer to the paper [105], where stability of higher-order terms was discussed and an alternative approach for more stable results was developed back in 2005. Yet, we have found that some programs don't have to give reliable results even for the zero order multipole coefficients (charges). On the other hand, when using the newest GDMA code, we have obtained convergence of coefficients with (probably) similar properties as in [105]. We have found that when the molecules are rotated and their multipole expansion in spherical tensor formalism is rotated back, we obtain the same coefficients for both utilized codes, which is certainly reassuring.

During this work, we have developed software, that can obtain coefficients of distributed multipole expansion calculated by integration of density matrix in Molpro DMA or GDMA program, quickly evaluate interaction of multipoles up to arbitrary order and rotate multipole expansion by an arbitrary angle. This software can also work with Mulliken population analysis obtained from Molpro property program, which gave much less meaningful results than the DMA approach, and thus we did not include any results into this text.

We have also tested DMA for molecular solids in the fragmented approach, where we have obtained good agreement with SCF energy for two body contributions for NH_3 and methanol. The CO_2 molecule cannot be described well by distributed approach by considering coefficients only up to third order of expansion and fourth order is needed. The fourth order then introduces the quadrupole-quadrupole interaction, which is essential for CO_2. This means, that the description of the CO_2 molecule relies on high order coefficients, which are possibly unstable with respect to the size of the basis set. Huge relative numerical inaccuracies in ab-initio energies for nonpolar methane molecule mean that DMA is probably more precise for description of electrostatic effects, but it also means that we do not have meaningful data for comparison.

We have noticed that the correlation energy is critically important in the

total energy. For most molecular solids, the SCF energy constitutes maximally 50 % of the whole binding energy (when comparing with MP2 calculations on dimers), reducing the applicability of this strategy only for contributions from really distant dimers. For methane, ammonia, and carbon dioxide, dispersion forces are considered essential [17]. One way to utilize DMA in accurate energy calculations can be to check if contributions from distant dimers are relevant. This might help to determine the cutoff distance for two body energies in future calculations of systems, where two body interactions are the most important. An interesting and unexpected result is that the method, which is used for obtaining the density matrix for the evaluation of multipole coefficients, does not influence the results much (comparison were done for SCF and CCSD(T)).

We have found out that DMA is not reliable in all instances. Therefore, refitting coefficients might be required. We tried to do so by using the classical fitting procedures in scipy library as well as Python bindings to the GSL library using the lsqfit package; however, this approach has failed due to performance issues with the Python programming language and properties of optimization algorithms in those codes. Subsequently, we have implemented multipole expansion in TensorFlow (as a Keras model), where we relied on automatic differentiations as well as on efficient optimizers, and we were able to fit dimer-dimer interactions in explicitly correlated CCSD(T) calculations. We have tried to develop a more general code, but when higher terms of multipole expansion were added, the fit completely failed. We assume that the space of initial parameters has to be scanned during the fitting procedure (for example by simple GridSearch algorithm), or the implementation was flawed. Unfortunately, when trying fitting in TensorFlow on a simple model of long-distance correlation energy, which inherently needs distributed treatment, we have not obtained good results (probably due to many existing local minima). This most likely means that the distributed nature of multipoles is really difficult to capture by fitting on benchmark data, but it is possible to develop a model describing electrostatic interaction in the spherical tensor formalism around only one site and fit it by applying techniques from machine learning.

This work can also be seen in the context of quickly growing capabilities of machine learning algorithms, efficient encodings of molecular structures and Δ-learning. Neural networks may be trained to predict coefficients of physical models describing contributions to intermolecular energy (multipole expansion, induction, dispersion). Some work has already been done in the area of prediction of multipole coefficients directly from the structure of molecules [83].

We would like to finish this text by pointing out that putting all the definitions and codes together was a rather daunting task. One has to start with the definition of spherical harmonics and Euler angles used in GDMA or Molpro DMA code, which relies on the definitions of Lebedev quadrature. This has to be combined with the definition in the paper [90], where the interaction of multipoles in the spherical tensor formalism is derived and with definitions in [106] and spherical package. All utilized codes were rather poorly documented. Going through the source code of the spherical package was needed to find the reference to the paper [106]. We also needed to check Molpro DMA and GDMA source codes to find out the units of calculated multipole coefficients. Elementary tests ensuring correctness of our implementation were made.

Bibliography

[1] Anthony Stone. The Theory of Intermolecular Forces. Oxford University Press, 01 2013.

[2] Andrew I. Cooper. Porous molecular solids and liquids. *ACS Central Science*, 3(6):544–553, Jun 2017.

[3] Niamh O'Reilly, Nicola Giri, and Stuart L James. Porous liquids. *Chemistry–A European Journal*, 13(11):3020–3025, 2007.

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

[5] Zheng Rong Chong, She Hern Bryan Yang, Ponnivalavan Babu, Praveen Linga, and Xiao-Sen Li. Review of natural gas hydrates as an energy resource: Prospects and challenges. *Applied energy*, 162:1633–1652, 2016.

[6] Gregory JO Beran. Modeling polymorphic molecular crystals with electronic structure theory. *Chemical reviews*, 116(9):5567–5613, 2016.

[7] Michael J. Gillan, Dario Alfè, and Angelos Michaelides. Perspective: How good is DFT for water? *The Journal of Chemical Physics*, 144(13):130901, 04 2016.

[8] Biswajit Santra, Jirí Klimeš, Dario Alfè, Alexandre Tkatchenko, Ben Slater, Angelos Michaelides, Roberto Car, and Matthias Scheffler. Hydrogen bonds and van der waals forces in ice at ambient and high pressures. *Phys. Rev. Lett.*, 107:185701, Oct 2011.

[9] Biswajit Santra, Jiří Klimeš, Alexandre Tkatchenko, Dario Alfè, Ben Slater, Angelos Michaelides, Roberto Car, and Matthias Scheffler. On the accuracy of van der Waals inclusive density-functional theory exchange-correlation functionals for ice at ambient and high pressures. *The Journal of Chemical Physics*, 139(15):154702, 10 2013.

[10] Markus Macher, Jirí Klimeš, Cesare Franchini, and Georg Kresse. The random phase approximation applied to ice. *The Journal of Chemical Physics*, 140(8), 2014.

[11] J Bauer, S Spanton, R Henry, J Quick, W Dziki, W Porter, and J Morris. Ritonavir: an extraordinary example of conformational polymorphism. *Pharm Res*, 18(6):859–866, June 2001.

[12] A Otero-De-La-Roza and Erin R Johnson. A benchmark for non-covalent interactions in solids. *The Journal of chemical physics*, 137(5), 2012.

[13] Anthony M Reilly and Alexandre Tkatchenko. Understanding the role of vibrations, exact exchange, and many-body van der waals interactions in the cohesive properties of molecular crystals. *The Journal of chemical physics*, 139(2), 2013.

[14] Ctirad Cervinka and Michal Fulem. State-of-the-art calculations of sublimation enthalpies for selected molecular crystals and their computational uncertainty. *Journal of chemical theory and computation*, 13(6):2840–2850, 2017.

[15] Leeor Kronik and Jeffrey B Neaton. Excited-state properties of molecular solids from first principles. *Annual review of physical chemistry*, 67:587–616, 2016.

[16] Urszula Góra, Rafał Podeszwa, Wojciech Cencek, and Krzysztof Szalewicz. Interaction energies of large clusters from many-body expansion. *The Journal of chemical physics*, 135(22), 2011.

[17] Jaroslav Hofierka and Jiří Klimeš. Binding energies of molecular solids from fragment and periodic approaches. *Electronic Structure*, 3(3):034010, 2021.

[18] Lars Goerigk. A comprehensive overview of the dft-d3 london-dispersion correction. *Non-covalent interactions in quantum chemistry and physics*, pages 195–219, 2017.

[19] Wojciech Jankiewicz, Rafał Podeszwa, and Henryk A Witek. Dispersion-corrected dft struggles with predicting three-body interaction energies. *Journal of Chemical Theory and Computation*, 14(10):5079–5089, 2018.

[20] Michael J Deible, Odbadrakh Tuguldur, and Kenneth D Jordan. Theoretical study of the binding energy of a methane molecule in a (h2o) 20 dodecahedral cage. *The Journal of Physical Chemistry B*, 118(28):8257–8263, 2014.

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

[22] Ke Liao and Andreas Grüneis. Communication: Finite size correction in periodic coupled cluster theory calculations of solids. *The Journal of Chemical Physics*, 145(14), 2016.

[23] Mauro Del Ben, Jürg Hutter, and Joost VandeVondele. Second-order møller–plesset perturbation theory in the condensed phase: An efficient and massively parallel gaussian and plane waves approach. *Journal of chemical theory and computation*, 8(11):4177–4188, 2012.

[24] Tobias Schäfer, Benjamin Ramberger, and Georg Kresse. Quartic scaling mp2 for solids: A highly parallelized algorithm in the plane wave basis. *The Journal of Chemical Physics*, 146(10), 2017.

[25] Marcin Modrzejewski, Sirous Yourdkhani, and Jiri Klimes. Random phase approximation applied to many-body noncovalent systems. *Journal of Chemical Theory and Computation*, 16(1):427–442, 2019.

[26] AJ Stone and M Alderton. Distributed multipole analysis methods and applications. *Molecular Physics*, 100(1):221–233, 2002.

[27] Eric R Scerri. Just how ab initio is ab initio quantum chemistry? *Foundations of Chemistry*, 6(1):93–116, 2004.

[28] William MC Foulkes, Lubos Mitas, RJ Needs, and Guna Rajagopal. Quantum monte carlo simulations of solids. *Reviews of Modern Physics*, 73(1):33, 2001.

[29] Priyanka Makkar and Narendra Nath Ghosh. A review on the use of dft for the prediction of the properties of nanomaterials. *RSC Adv.*, 11:27897–27924, 2021.

[30] Alyssa J. R. Hensley, Kushal Ghale, Carolin Rieg, Thanh Dang, Emily Anderst, Felix Studt, Charles T. Campbell, Jean-Sabin McEwen, and Ye Xu. Dft-based method for more accurate adsorption energies: An adaptive sum of energies from rpbe and vdw density functionals. *The Journal of Physical Chemistry C*, 121(9):4937–4945, Mar 2017.

[31] Kieron Burke, Jan Werschnik, and E. K. U. Gross. Time-dependent density functional theory: Past, present, and future. *The Journal of Chemical Physics*, 123(6):062206, 08 2005.

[32] Frederico D. Novaes, Antônio J. R. da Silva, and A. Fazzio. Density functional theory method for non-equilibrium charge transport calculations: Transampa. *Brazilian Journal of Physics*, 36(3a):799–807, Sep 2006.

[33] Peter Ring. *Concept of covariant density functional theory*, pages 1–20. 03 2016.

[34] Junjie Yang and J. Piekarewicz. Covariant density functional theory in nuclear physics and astrophysics. *Annual Review of Nuclear and Particle Science*, 70(1):21–41, 2020.

[35] James D Whitfield, Jun Yang, Weishi Wang, Joshuah T Heath, and Brent Harrison. Quantum computing 2022. *arXiv preprint arXiv:2201.09877*, 2022.

[36] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung,

Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, Oct 2019.

[37] Feng Pan, Keyang Chen, and Pan Zhang. Solving the sampling problem of the sycamore quantum circuits. *Phys. Rev. Lett.*, 129:090502, Aug 2022.

[38] Ruslan N. Tazhigulov, Shi-Ning Sun, Reza Haghshenas, Huanchen Zhai, Adrian T.K. Tan, Nicholas C. Rubin, Ryan Babbush, Austin J. Minnich, and Garnet Kin-Lic Chan. Simulating models of challenging correlated molecules and materials on the sycamore quantum processor. *PRX Quantum*, 3:040318, Nov 2022.

[39] Hongye Yu and Tzu-Chieh Wei. Quantum zeno approach for molecular energies with maximum commuting initial hamiltonians. *Phys. Rev. Res.*, 3:013104, Feb 2021.

[40] Hongye Yu, Deyu Lu, Qin Wu, and Tzu-Chieh Wei. Geometric quantum adiabatic methods for quantum chemistry. *Phys. Rev. Res.*, 4:033045, Jul 2022.

[41] V. C. Aguilera-Navarro, G. A. Estévez, and R. Guardiola. Variational and perturbative schemes for a spiked harmonic oscillator. *Journal of Mathematical Physics*, 31(1):99–104, 01 1990.

[42] Jong H Choi, Charles F Lebeda, and Richard P Messmer. Variational principle for excited states: Exact formulation and other extensions. *Chemical Physics Letters*, 5(8):503–506, 1970.

[43] Hong-Zhou Ye, Matthew Welborn, Nathan D Ricke, and Troy Van Voorhis. σ-scf: A direct energy-targeting method to mean-field excited states. *The Journal of chemical physics*, 147(21), 2017.

[44] Jacqueline AR Shea, Elise Gwin, and Eric Neuscamman. A generalized variational principle with applications to excited state mean field theory. *Journal of chemical theory and computation*, 16(3):1526–1540, 2020.

[45] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[46] Naoki Nakatani. Matrix product states and density matrix renormalization group algorithm. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier, 2018.

[47] Frank Jensen. *Introduction to computational chemistry*. John Wiley & sons, 2007.

[48] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Dover Publications, 1996.

[49] Andreas Gruneis. *Diagramatic techniques for extended systems: MP2 and CCSD*. PhD thesis, University of Vienna, 2011.

[50] Tracy P Hamilton and Peter Pulay. Direct inversion in the iterative subspace (diis) optimization of open-shell, excited-state, and small multiconfiguration scf wave functions. *The Journal of chemical physics*, 84(10):5728–5734, 1986.

[51] Christopher Seidl and Giuseppe MJ Barca. Q-next: A fast, parallel, and diagonalization-free alternative to direct inversion of the iterative subspace. *Journal of Chemical Theory and Computation*, 18(7):4164–4176, 2022.

[52] Werner Dobrautz. *Development of Full Configuration Interaction Quantum Monte Carlo Methods for Strongly Correlated Electron Systems*. PhD thesis, Institut fur Theoretische Chemie der Universitat Stuttgart, 2019. Available at `https://d-nb.info/1197056459/34`.

[53] Christopher R Myers, Cyrus J Umrigar, James P Sethna, and John D Morgan III. Fock's expansion, kato's cusp conditions, and the exponential ansatz. *Physical Review A*, 44(9):5537, 1991.

[54] Dieter Cremer. Møller–plesset perturbation theory: From small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(4):509–530, 2011.

[55] F. Coester. Bound states of a many-particle system. *Nuclear Physics*, 7:421–424, 1958.

[56] F. Coester and H. Kümmel. Short-range correlations in nuclear wave functions. *Nuclear Physics*, 17:477–485, 1960.

[57] Jiří Čížek. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 05 2004.

[58] J. Čižek and J. Paldus. Correlation problems in atomic and molecular systems iii. rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methodst. *International Journal of Quantum Chemistry*, 5(4):359–379, 1971.

[59] Heiko Hergert. A guided tour of ab initio nuclear many-body theory. *Frontiers in Physics*, 8, 2020.

[60] Rodney J Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291, 2007.

[61] Antoine Marie, Fábris Kossoski, and Pierre-François Loos. Variational coupled cluster for ground and excited states. *The Journal of Chemical Physics*, 155(10), 2021.

[62] Marcin Modrzejewski, Sirous Yourdkhani, Szymon Smiga, and Jiri Klimes. Random-phase approximation in many-body noncovalent systems: Methane in a dodecahedral water cage. *Journal of Chemical Theory and Computation*, 17(2):804–817, 2021.

[63] John F. Stanton. Why ccsd(t) works: a different perspective. *Chemical Physics Letters*, 281(1):130–134, 1997.

[64] Judith Harl. *The linear response function in density functional theory: Optical spectra and improved description of the electron correlation.* PhD thesis, University of Vienna, 2008.

[65] Mark E Casida. Time-dependent density functional response theory of molecular systems: theory, computational methods, and functionals. *Theoretical and Computational Chemistry*, 4:391–439, 1996.

[66] Giampaolo Co'. Introducing the random phase approximation theory. *Universe*, 9(3):141, 2023.

[67] Agisilaos Chantzis, Adèle D Laurent, Carlo Adamo, and Denis Jacquemin. Is the tamm-dancoff approximation reliable for the calculation of absorption and fluorescence band shapes? *Journal of chemical theory and computation*, 9(10):4517–4525, 2013.

[68] FJ Dyson, Marc Ross, Edwin E Salpeter, Silvan S Schweber, MK Sundaresan, WM Visscher, and Hans A Bethe. Meson-nucleon scattering in the tamm-dancoff approximation. *Physical Review*, 95(6):1644, 1954.

[69] Thom H Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of chemical physics*, 90(2):1007–1023, 1989.

[70] David E Woon and Thom H Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon. *The Journal of chemical physics*, 98(2):1358–1371, 1993.

[71] Asger Halkier, Trygve Helgaker, Poul Jørgensen, Wim Klopper, and Jeppe Olsen. Basis-set convergence of the energy in molecular hartree–fock calculations. *Chemical Physics Letters*, 302(5-6):437–446, 1999.

[72] Huub JJ Van Dam, Joop H Van Lenthe, and Peter Pulay. The size consistency of multi-reference møller–plesset perturbation theory. *Molecular Physics*, 93(3):431–439, 1998.

[73] Bernd Doser, Jan Zienau, Lucien Clin, Daniel S Lambrecht, and Christian Ochsenfeld. A linear-scaling mp2 method for large molecules by rigorous integral-screening criteria. *Zeitschrift für Physikalische Chemie*, 224(3-4):397–412, 2010.

[74] Jan Almlöf. Elimination of energy denominators in møller—plesset perturbation theory by a laplace transform approach. *Chemical physics letters*, 181(4):319–320, 1991.

[75] Merzuk Kaltak, Jiri Klimes, and Georg Kresse. Low scaling algorithms for the random phase approximation: Imaginary time and laplace transformations. *Journal of chemical theory and computation*, 10(6):2498–2507, 2014.

[76] Honghui Shang and Jinlong Yang. Implementation of laplace transformed mp2 for periodic systems with numerical atomic orbitals. *Frontiers in chemistry*, 8:589992, 2020.

[77] Seiichiro Ten-no. Initiation of explicitly correlated slater-type geminal theory. *Chemical Physics Letters*, 398(1-3):56–61, 2004.

[78] Thomas Schraivogel, Aron J Cohen, Ali Alavi, and Daniel Kats. Transcorrelated coupled cluster methods. *The Journal of Chemical Physics*, 155(19), 2021.

[79] Thomas Schraivogel, Evelin Martine Christlmaier, Pablo López Ríos, Ali Alavi, and Daniel Kats. Transcorrelated coupled cluster methods. ii. molecular systems. *The Journal of Chemical Physics*, 158(21), 2023.

[80] Hans-Joachim Werner, Peter J. Knowles, Frederick R. Manby, Joshua A. Black, Klaus Doll, Andreas Heßelmann, Daniel Kats, Andreas Köhn, Tatiana Korona, David A. Kreplin, Qianli Ma, III Miller, Thomas F., Alexander Mitrushchenkov, Kirk A. Peterson, Iakov Polyak, Guntram Rauhut, and Marat Sibaev. The Molpro quantum chemistry package. *The Journal of Chemical Physics*, 152(14):144107, 04 2020.

[81] Katarzyna Madajczyk, Piotr S Zuchowski, Filip Brzek, Lukasz Rajchel, Dariusz Kedziera, Marcin Modrzejewski, and Michal Hapka. Dataset of noncovalent intermolecular interaction energy curves for 24 small high-spin open-shell dimers. *The Journal of Chemical Physics*, 154(13), 2021.

[82] Celeste Sagui, Christopher Roland, Lee G Pedersen, and Thomas A Darden. New distributed multipole metdhods for accurate electrostatics in large-scale biomolecular simulations. *New Algorithms for Macromolecular Simulation*, pages 297–312, 2006.

[83] Moritz Thurlemann, Lennard Boselt, and Sereina Riniker. Learning atomic multipoles: prediction of the electrostatic potential with equivariant graph neural networks. *Journal of Chemical Theory and Computation*, 18(3):1701–1710, 2022.

[84] B Wang, W Zhang, and W Cai. Fast multipole method for 3-d laplace equation in layered media, submitted to comput. *Phys. Commun*, 2019.

[85] Nail A Gumerov and Ramani Duraiswami. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier, 2005.

[86] Eric Darve. The fast multipole method: numerical implementation. *Journal of Computational Physics*, 160(1):195–240, 2000.

[87] Anna-Karin Tornberg and Leslie Greengard. A fast multipole method for the three-dimensional stokes equations. *Journal of Computational Physics*, 227(3):1613–1619, 2008.

[88] Johannes Tausch. The fast multipole method for arbitrary green's functions. 329, 01 2003.

[89] Andreas Ross. Multipole expansion at the level of the action. *Physical Review D*, 85(12):125033, 2012.

[90] C. G. Gray. On the theory of multipole interactions. *Canadian Journal of Physics*, 46(2):135–139, 1968.

[91] Wu-Ki Tung. *Group theory in physics*, volume 1. World Scientific, 1985.

[92] Dmitriĭ Aleksandrovich Varshalovich, Anatolï Nikolaevitch Moskalev, and Valerij Kel'manovĭc Khersonskii. *Quantum theory of angular momentum*. World Scientific, 1988.

[93] B. C. Carlson and G. S. Rushbrooke. On the expansion of a coulomb potential in spherical harmonics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 46(4):626–633, 1950.

[94] Morris Edgar Rose. *Elementary theory of angular momentum*. Courier Corporation, 1995.

[95] AP Yutsis, IB Levinson, and VV Vanagas. The theory of angular momentum (israel program for scientific translations, jerusalem, 1962). 25.

[96] SL Price, AJ Stone, and M Alderton. Explicit formulae for the electrostatic energy, forces and torques between a pair of molecules of arbitrary symmetry. *Molecular Physics*, 52(4):987–1001, 1984.

[97] AJ Stone and RJA Tough. Spherical tensor theory of long-range intermolecular forces. *Chemical physics letters*, 110(2):123–129, 1984.

[98] Mike Boyle. moble/spherical: Release v1.0.13, February 2023.

[99] Rick Beatson, Leslie Greengard, et al. A short course on fast multipole methods. *Wavelets, multilevel methods and elliptic PDEs*, 1:1–37, 1997.

[100] Radovan Bast. Numgrid: Numerical integration grid for molecules, 1 2021.

[101] John Burkardt. Sphere_lebedev_rule: Quadrature rules for the unit sphere, 2010.

[102] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[103] Daniel G. A. Smith, Lori A. Burns, Andrew C. Simmonett, Robert M. Parrish, Matthew C. Schieber, Raimondas Galvelis, Peter Kraus, Holger Kruse, Roberto Di Remigio, Asem Alenaizan, Andrew M. James, Susi Lehtola, Jonathon P. Misiewicz, Maximilian Scheurer, Robert A. Shaw, Jeffrey B. Schriber, Yi Xie, Zachary L. Glick, Dominic A. Sirianni, Joseph Senan O'Brien, Jonathan M. Waldrop, Ashutosh Kumar, Edward G. Hohenstein, Benjamin P. Pritchard, Bernard R. Brooks, III Schaefer, Henry F., Alexander Yu. Sokolov, Konrad Patkowski, III DePrince, A. Eugene, Uğur Bozkaya, Rollin A. King, Francesco A. Evangelista, Justin M. Turney, T. Daniel Crawford, and C. David Sherrill. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *The Journal of Chemical Physics*, 152(18):184108, 05 2020.

[104] Jiri Klimes and Jaroslav Hofierka. Dataset for "Binding energies of molecular solids from fragment and periodic approaches", October 2021.

[105] Anthony J Stone. Distributed multipole analysis: Stability for large basis sets. *Journal of Chemical Theory and Computation*, 1(6):1128–1132, 2005.

[106] Nail A Gumerov and Ramani Duraiswami. Recursive computation of spherical harmonic rotation coefficients of large degree. In *Excursions in Harmonic Analysis, Volume 3: The February Fourier Talks at the Norbert Wiener Center*, pages 105–141. Springer, 2015.

List of Figures

List of Tables

List of Abbreviations

A. Efficient and precise numerical evaluation of Wigner D matrices and spherical harmonics

We use the `spherical` package [98], implementing the recursive algorithm developed in [106]. This approach should guarantee reasonable precision up to $l \lesssim 10^3$ which certainly is enough for our usage. We will show the basic ideas behind this approach, for details and considerations regarding stability, consult [106].

By using matrices

$$A(\gamma) = \begin{pmatrix} \sin\gamma & \cos\gamma & 0 \\ -\cos\gamma & \sin\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B(\beta) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\cos\beta & \sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix}, \tag{A.1}$$

we may define an arbitrary rotation matrix

$$Q(\alpha, \beta, \gamma) = A(\gamma)B(\beta)A^T(\alpha)\,. \tag{A.2}$$

Furthermore, we may rewrite any function as[1]

$$\begin{aligned} f(\Theta, \varphi) &= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_n^m Y_n^m(\Theta, \varphi) = \\ &= \sum_{n=0}^{\infty} \sum_{m'=n}^{n} \hat{C}_n^{m'} Y_n^{m'}(\hat{\Theta}, \hat{\varphi}) = \hat{f}(\hat{\Theta}, \hat{\varphi})\,, \end{aligned} \tag{A.3}$$

where the `hat` symbol means that the symbol is evaluated in the frame of reference rotated by the $Q(\alpha, \beta, \gamma)$ matrix. Putting $C_n^m = \delta_{nm}$ we obtain from (A.3)

$$Y_n^\nu(\Theta, \varphi) = e^{i\nu\alpha} \sum_{m'=-n}^{n} H_n^{m'\nu}(\beta) e^{-im'\gamma} Y_n^{m'}(\hat{\Theta}, \hat{\varphi})\,. \tag{A.4}$$

Explicit expression for $H_n^{m',m}$ is acquired from Wigner formula

$$\begin{aligned} H_n^{m',m}(\beta) &= \varepsilon_{m'}\varepsilon_m \rho_n^{m',m} \sum_{\sigma=\max(0,-m'-m)}^{\min(n-m',n-m)} (-1)^{n-\sigma} h_n^{m',m,\sigma}(\beta) \\ &= \varepsilon_{m'}\varepsilon_m \rho_n^{m',m} \sum_{\sigma=-\infty}^{\infty} (-1)^{n-\sigma} h_n^{m',m,\sigma}(\beta)\,, \end{aligned} \tag{A.5}$$

where the definition $\frac{1}{(-n)!} = 0$ for $n = 1, 2, \ldots$ leading to the expression on the last line is motivated by the limit of $\frac{1}{\Gamma(-n)}$ for $n \to 1, 2, \ldots$ (see equation (A.6)). Furthermore, ε_m is $(-1)^m$ if $m \in \mathbb{Z}$ or 1 otherwise and $\rho_n^{m'm} =$

[1] Please note that in this appendix, we use notation and definitions from the paper [106] rather than the one used in the rest of this text. We have made this inconsistent choice to make going through the paper [106] easier. The small Wigner d matrices are related to $H_n^{m'\nu}$ and are written with flipped indices in [106]. The paper [106] uses passive interpretation of Euler angles in the z–y–z convention. In this convention, the entities $H_n^{m',\nu}$ are real and the relation $d_n^{m'm} = \varepsilon_{m'}\varepsilon_{-m} H_n^{m'm}$ holds.

$\sqrt{(n+m)!(n-m)!(n+m')!(n-m')!}$. There also exists identity (which would in principle allow us to evaluate spherical harmonics and Wigner D matrices)

$$h_n^{m',m,\sigma}(\beta) = \frac{\cos^{2\sigma+m+m'}\frac{\beta}{2}\sin^{2n-2\sigma-m-m'}\frac{\beta}{2}}{\sigma!(n-m'-\sigma)!(n-m-\sigma)!(m'+m+\sigma)!}\,. \tag{A.6}$$

Yet, there is a better way to do so. Defining $\mathbf{Rot}_n(Q(\alpha,\beta,\gamma))$ as invariant subspace of $e^{im\alpha}H_n^{m'm}e^{-im'\gamma}$ (see equation (19) and discussion above in [106]) we have matrix elements of $\mathbf{Rot}_n(Q_y(\alpha))$ given by (where the $(-1)^{m'}$ accounts for the definitions of Euler angles in the paper [106] and the fact that $H_n^{m'm}$ are not small Wigner d matrices)

$$(-1)^{m'}H_n^{m'm}(\beta)\,. \tag{A.7}$$

The commutativity of rotations around the y axis

$$Q_y(\beta)Q_y'(0) = Q_y'(0)Q_y(\beta)\,,\quad Q_y'(\beta) = \left.\frac{\mathrm{d}Q_y(x)}{\mathrm{d}x}\right|_{x=\beta} \tag{A.8}$$

gives us a relation for the n-th invariant rotation transform subspace

$$\mathbf{Rot}_nQ_y(\beta)\mathbf{Rot}_nQ_y'(0) = \mathbf{Rot}_nQ_y'(0)\mathbf{Rot}_nQ_y(\beta)\,. \tag{A.9}$$

Taking the derivative of the right-hand side of equation (A.5) with respect to β and evaluating it at $\beta = 0$ yields

$$\left.\frac{\mathrm{d}H_n^{m',m}(\beta)}{\mathrm{d}\beta}\right|_{\beta=0} = c_n^{m'-1}\delta_{m,m'-1} + c_n^{m'}\delta_{m,m'+1}\,, \tag{A.10}$$

where

$$c_n^m = \frac{(-1)^m\mathrm{sgn}(m)}{2}\sqrt{(n-m)(n+m+1)}\,,\quad m\in\{-n-1,\dots,n\}\,. \tag{A.11}$$

Applying the relation (A.7) to (A.9) gives us

$$\sum_{\nu=-n}^{n}H_n^{m',\nu}(\beta)(-1)^{\nu}\left.\frac{\mathrm{d}H_n^{\nu,m}(\beta)}{\mathrm{d}\beta}\right|_{\beta=0} = \sum_{\nu=-n}^{n}\left.\frac{\mathrm{d}H_n^{m',\nu}(\beta)}{\mathrm{d}\beta}\right|_{\beta=0}(-1)^{\nu}H_n^{\nu,m}(\beta)\,. \tag{A.12}$$

By the virtue of equation (A.10), we obtain (see equation (50) in [106])

$$d_n^{m-1}H_n^{m',m-1} - d_n^mH_n^{m',m+1} = d_n^{m'-1}H_n^{m'-1,m} - d_n^{m'}H_n^{m'+1,m}\,, \tag{A.13}$$

using the definition

$$d_n^m = \frac{\mathrm{sgn}(m)}{2}\sqrt{(n-m)(n+m+1)}\,. \tag{A.14}$$

The importance of the relation (A.13) (which is one of the main results of the rather long paper [106]) is that it uses only values within the same subspace n. This is in direct contrast with other methods. By using easily calculable values of $H_n^{m,m'}$ (for example by setting $n = m$ where only $\sigma = 0$ contributes in (A.6)) or values from $n-1$ subspace, one may obtain coefficients for all possible indices. Wigner D matrices and (spin-weighted) spherical harmonics can be efficiently and precisely evaluated using those. Even though we do not use particularly high order of spherical harmonics, the spherical package provides values of spherical harmonics for all possible indices of order less than some l from only one calculation, making it the main reason for utilizing this library.

B. Code description

Our code is available at GitHub repository[1]. We require that packages quaternionic, spherical, numpy, typing and regex are installed. The "main" file is `lib.py` where all algorithms important for any user are implemented. A user can use static methods in the DMA class to construct an array of entries (corresponding to sites in DMA) which can be then used to evaluate intermolecular energy. The Entry structure holds `position`, `l` (order of the expansion) and `data` (actual coefficients of spherical harmonic expansion). Our code can read data from Molpro (everybody is discouraged from using those results) and GDMA. Both codes return total multipole at origin calculated by equation (3.47) which is ignored by our algorithms. This site does not have any effect on interaction whatsoever, on the other hand, it can be used to check that the results are meaningful and in expected units. Note that the multipoles Q_{lm} should be obtained by equations (3.50).

The code also implements Mulliken population analysis, which was studied in the beginning without obtaining any satisfactory results.

We hope that when input in invalid format is passed to our code, the program safely exits with an exception and does not continue with meaningless or corrupted input values.

Psi4 + GDMA

To obtain distributed multipole expansion from GDMA, use the function `get_from_gdma_file`. To obtain coefficients of multipole expansion, one has to first calculate the wave function and save it into fchk file. This can be done by code (B.1). The next step is to produce a command file for the GDMA code (see code (B.2)), where data will be saved to a file called dma4.punch.

Listing B.1: Psi4 command for saving wave function in fchk file.

```
energy, wfn = energy('ccsd(t)', return_wfn=True)
fchk(wfn,'output.fchk')
```

Listing B.2: GDMA command.

```
File output.fchk
Angstrom
Multipoles
  switch 10
  Limit 10
  Punch dma4.punch
Start
Finish
```

One issue with Psi4 calculation is that Psi4 automatically translates and rotates geometry. This can be prevented by writing the code (B.3) directly into the geometry specification (but before any geometry is actually defined).

[1] `https://github.com/amendl/bc_final_code`

Listing B.3: Commans for preventing rotation and translation in Psi4 package.

```
nocom
noreorient
```

Molpro

To read data from Molpro output file, use the static function `get_from_molpro_file` in the DMA class. Molpro uses a rather strange way to represent IDs of calculated wave functions, and the Molpro DMA code ignores which density matrix was given to the property program. Therefore, the density matrix has to be explicitly saved and loaded. Be careful, however, since you can overwrite results from other calculations. The command (B.4) was tested and works flawlessly.

Listing B.4: Molpro command.

```
{ccsd(t);DM,7000.2}
{dma;density,7000.2;limit,,5}
```

TensorFlow models

In listing (B.5), we present a simplified code for fitting electrostatic and correlation energies using the TensorFlow library. The precalculated spherical harmonics are obtained as in (B.6).

Listing B.5: Simplified TensorFlow model for dipole-dipole and correlation interaction.

```
class CorrelationLayer(keras.layers.Layer):
    def __init__(self, initializer = keras.initializers.
        GlorotUniform()):
        super(CorrelationLayer, self).__init__()
        self.initializer = initializer
    def build(self, input_shape):
        self.W = self.add_weight(shape=(36,),initializer=self.
            initializer, trainable=True, name='
            correlation_coefficients')
    def call(self, inputs,**kwargs):
        return tf.reduce_sum(tf.multiply(inputs, self.W),axis=-1)
class DipoleLayer(keras.layers.Layer):
    def __init__(self, initializer = keras.initializers.
        GlorotUniform()):
        super(DipoleLayer, self).__init__()
        self.initializer = initializer
    def build(self, input_shape):
        self.Q0 = self.add_weight(shape=(1,),initializer=self.
            initializer,trainable=True, name="Q0")
        self.Qc = self.add_weight(shape=(1,),initializer=self.
            initializer,trainable=True, name="Qc")
        self.Qd = self.add_weight(shape=(1,),initializer=self.
            initializer,trainable=True, name="Qd")
    def call(self, inputs,**kwargs):
        R = tf.concat([-self.Qc,self.Q0,self.Qd],axis=0)
```

```
        I = tf.concat([-self.Qd,self.Q0,self.Qc],axis=0)
        r,i = inputs
        a = tf.complex(R,I)
        return tf.math.real(tf.reduce_sum(tf.math.multiply(tf.
            tensordot(a,a,axes=0), tf.complex(r,i)),axis=[-1,-2]))
def architecture():
    i1 = keras.layers.Input(shape=(36,))
    i2 = keras.layers.Input(shape=(3,3))
    i3 = keras.layers.Input(shape=(3,3))
    correlationLayer = CorrelationLayer()
    dipoleLayer = DipoleLayer()
    x1 = correlationLayer(i1)
    x2 = dipoleLayer([i2,i3])
    added = keras.layers.Add()([x1, x2])
    return (keras.Model(inputs=[i1,i2,i3], outputs=added),
        correlationLayer,dipoleLayer)
model = architecture()
model.compile(optimizer='adam',loss='mean_squared_error')

# loading data, transforming data, code omitted

model.fit(x = [distances_correlation_fit,np.real(distances_dipole
    ),np.imag(distances_dipole)],y=energies,epochs=100,batch_size
    =16,shuffle=True)
```

Listing B.6: Creation of the dataset of precalculated spherical harmonics for dipole-dipole interaction fitting in TensorFlow and GSL/lsqfit.

```
# loading files, energies, angles and iteration through data --
    code ommited

def bracket_factor(l1,l2):
    return math.sqrt(float(factorial(2*l1+2*l2 + 1))/(float(
        factorial(2*l1))*float(factorial(2*l2))))

R_angle     = quaternionic.array.from_spherical_coordinates(theta
    ,phi)
wigner      = spherical.Wigner(depth)
Y           = wigner.sYlm(0,R_angle)
D           = wigner.D(R_angle)
l1_index = 1
l2_index = 1

for m1_ in range(2*l1_index+1):
    m1=m1_-l1_index
    for m2_ in range(2*l2_index+1):
        m2=m2_-l2_index
        for m_ in range(2*(l1_index+l2_index)+1):
            m =m_ - l1_index - l2_index
            part = D[wigner.Dindex(l1_index+l2_index,m,0)]
            retval[m1_,m2_]+= (pow(-1.,l2_index)/pow(R,l1_index +
                l2_index + 1))*bracket_factor(l1_index,l2_index)*
               part*spherical.Wigner3j(l1_index,l2_index,depth,m1
               ,m2,m)
```

C. Supporting materials

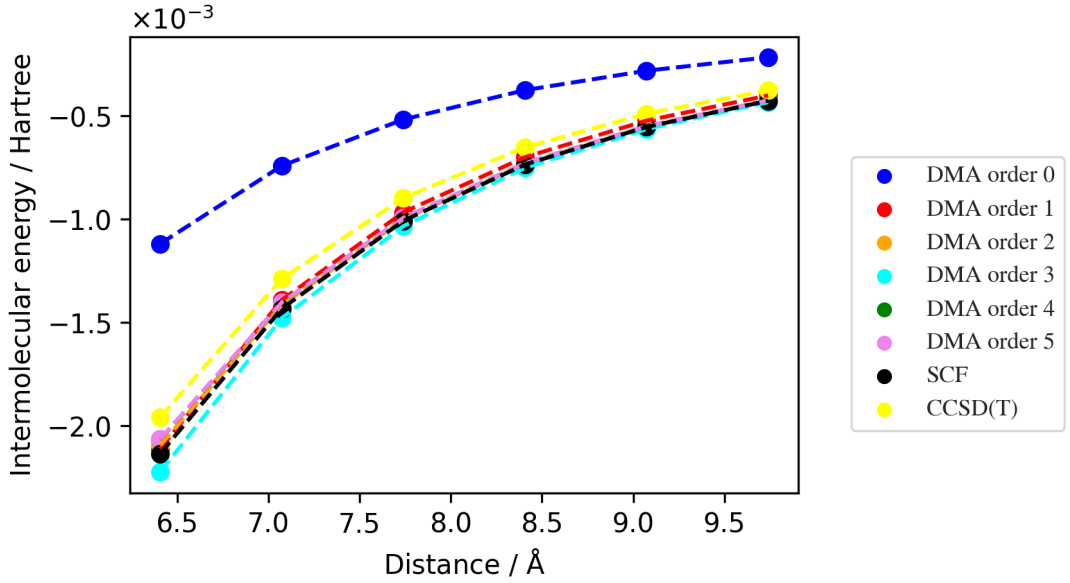
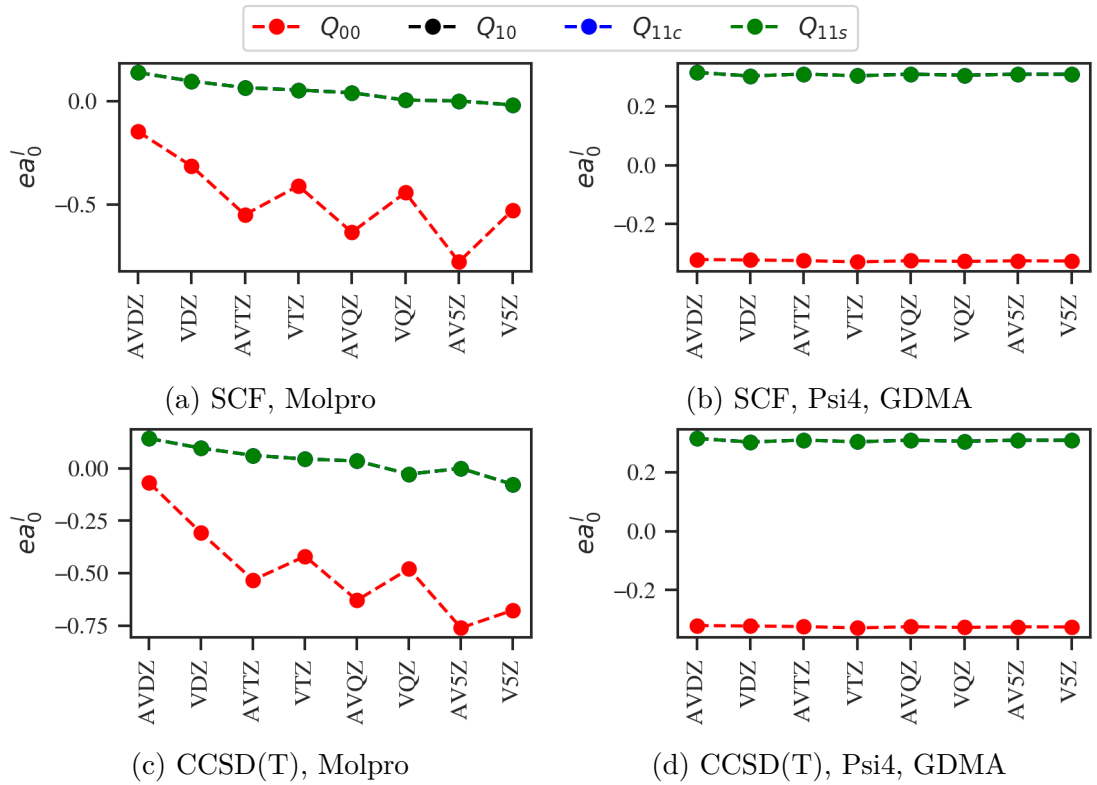


Figure C.1: Hydrogen fluoride binding curves – DMA of different orders with coefficients obtained from SCF calculation in Psi4 and GDMA on aug-cc-pV5Z basis, compared with SCF and CCSD(T) calculations from Molpro (BSIE correction from aug-cc-pVQZ and aug-cc-pV5Z basis). The distance between F atoms is shown.



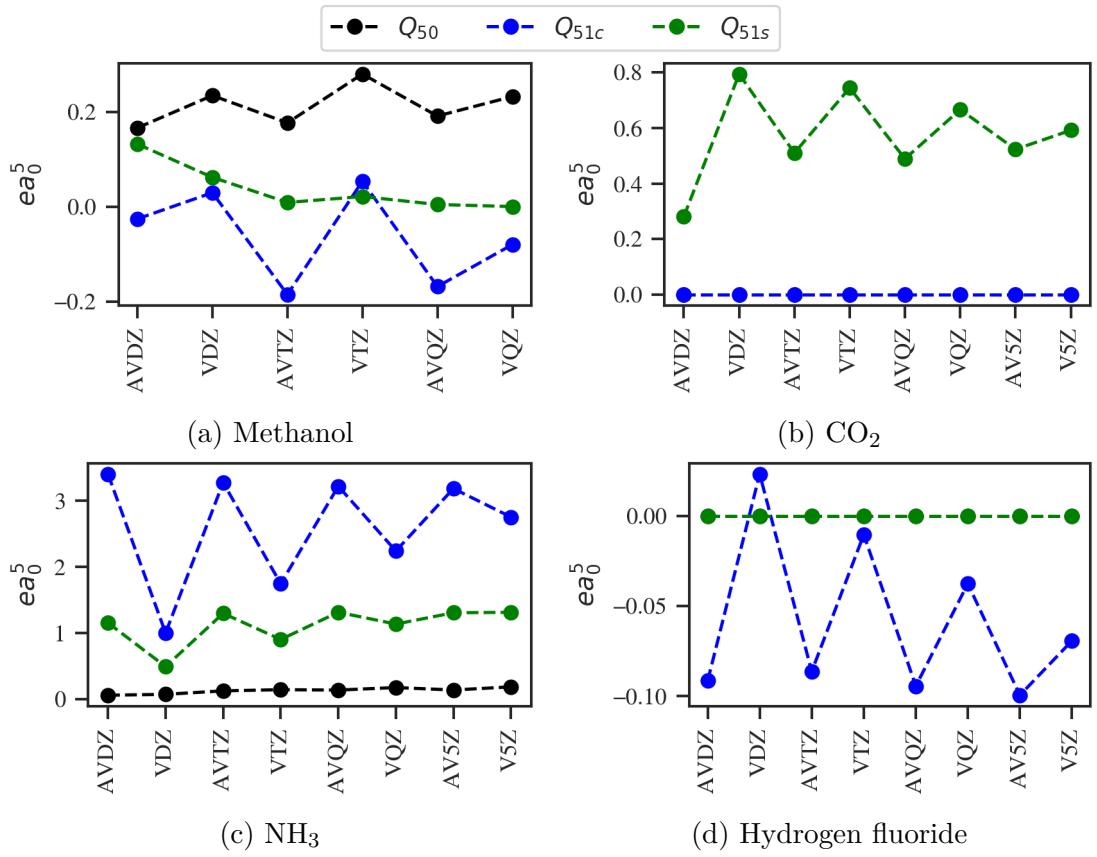
(a) SCF, Molpro

(b) SCF, Psi4, GDMA

(c) CCSD(T), Molpro

(d) CCSD(T), Psi4, GDMA

Figure C.2: NH_3: effects of basis, DMA code, and ab-initio method in the zeroth and first order of expansion. The y-axis units depend on coefficient order.



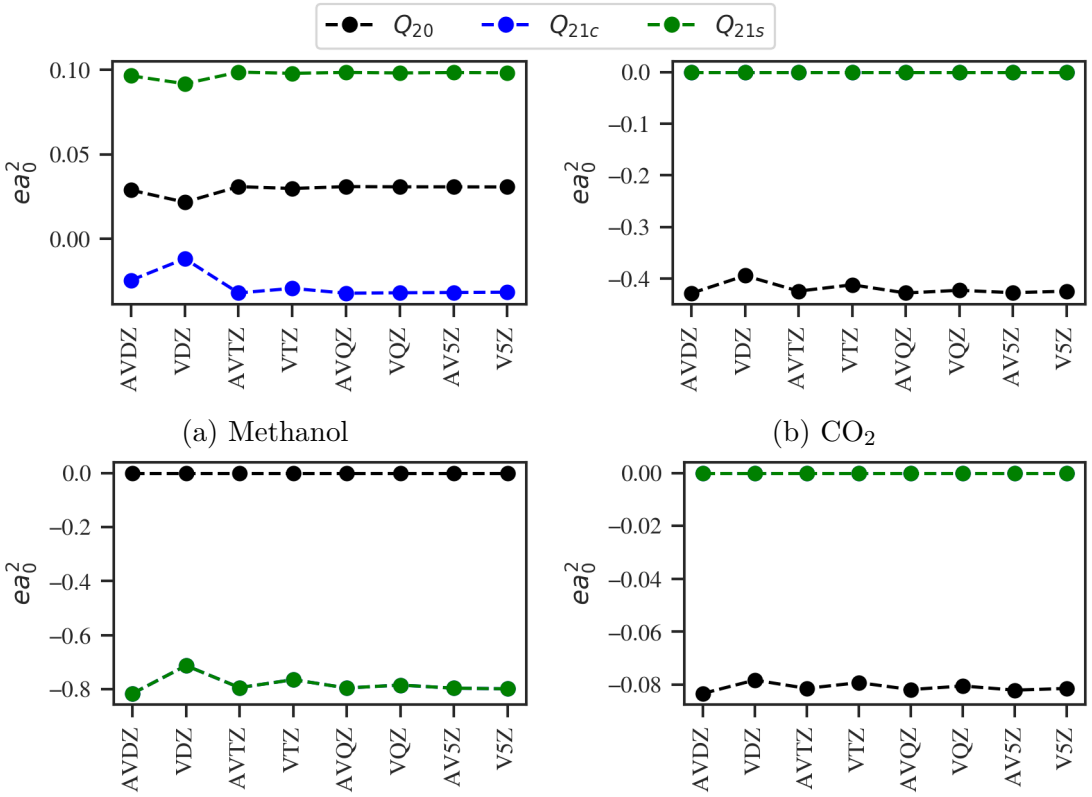
(a) Methanol

(b) CO_2

(c) NH_3

(d) Hydrogen fluoride

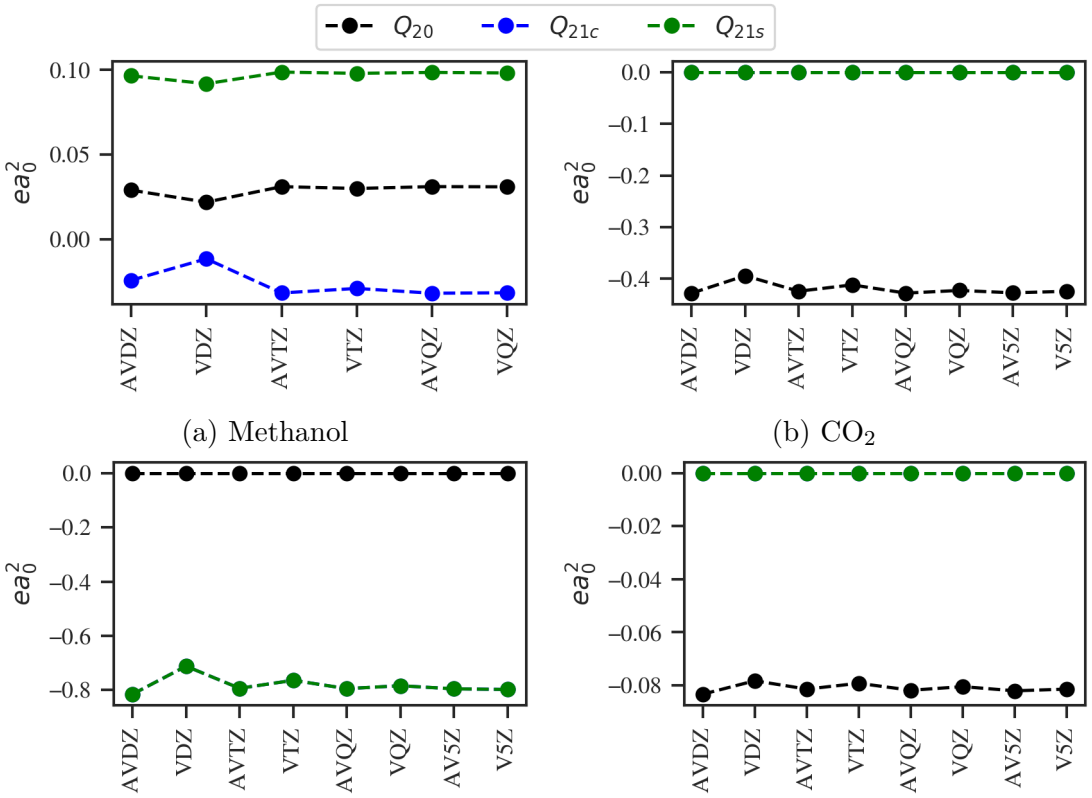
Figure C.3: Multipoles at 5th order from CCSD(T) density matrix and GDMA.



(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

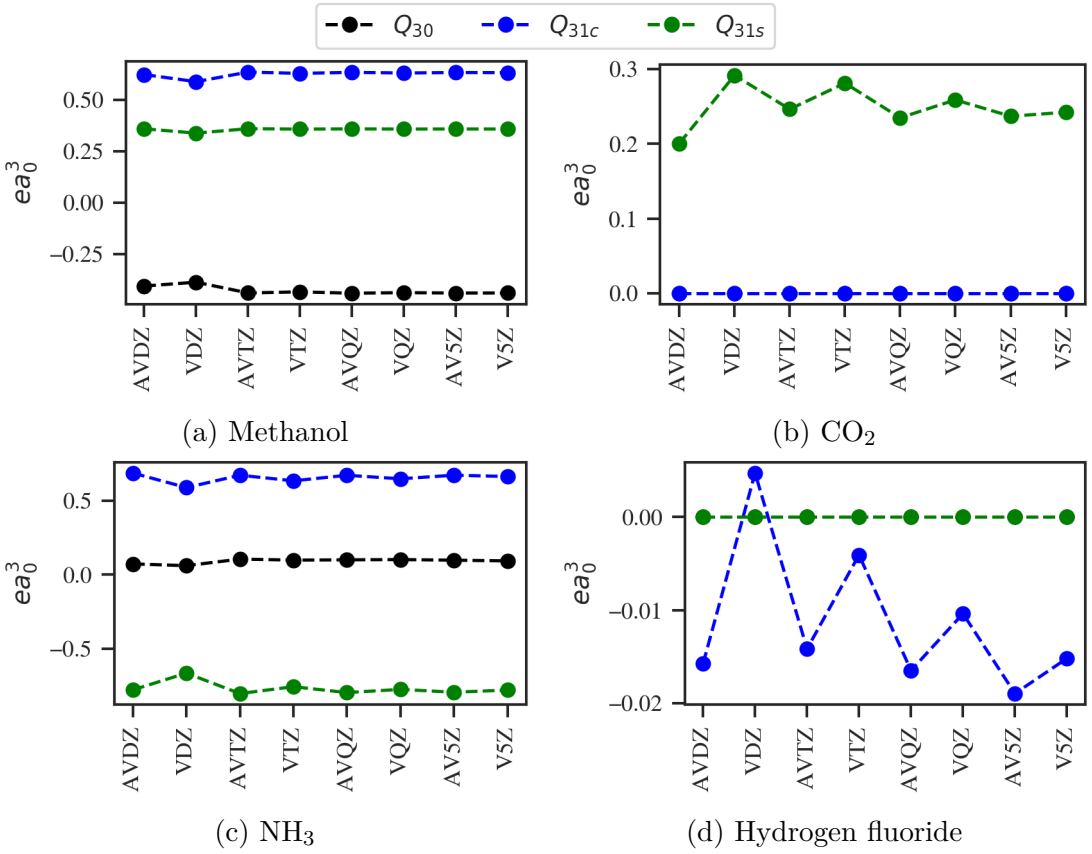
Figure C.4: Multipoles at 2nd order from SCF density matrix



(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

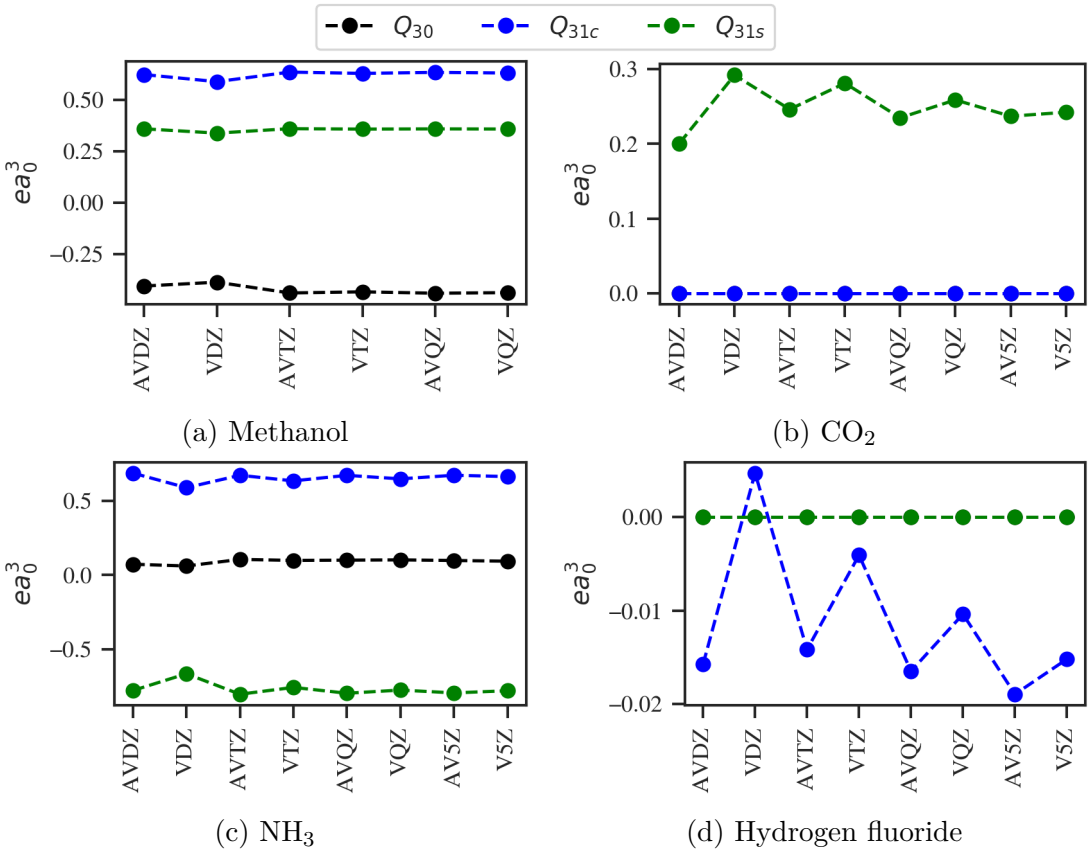
Figure C.5: Multipoles at 2nd order from CCSD(T) density matrix



(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

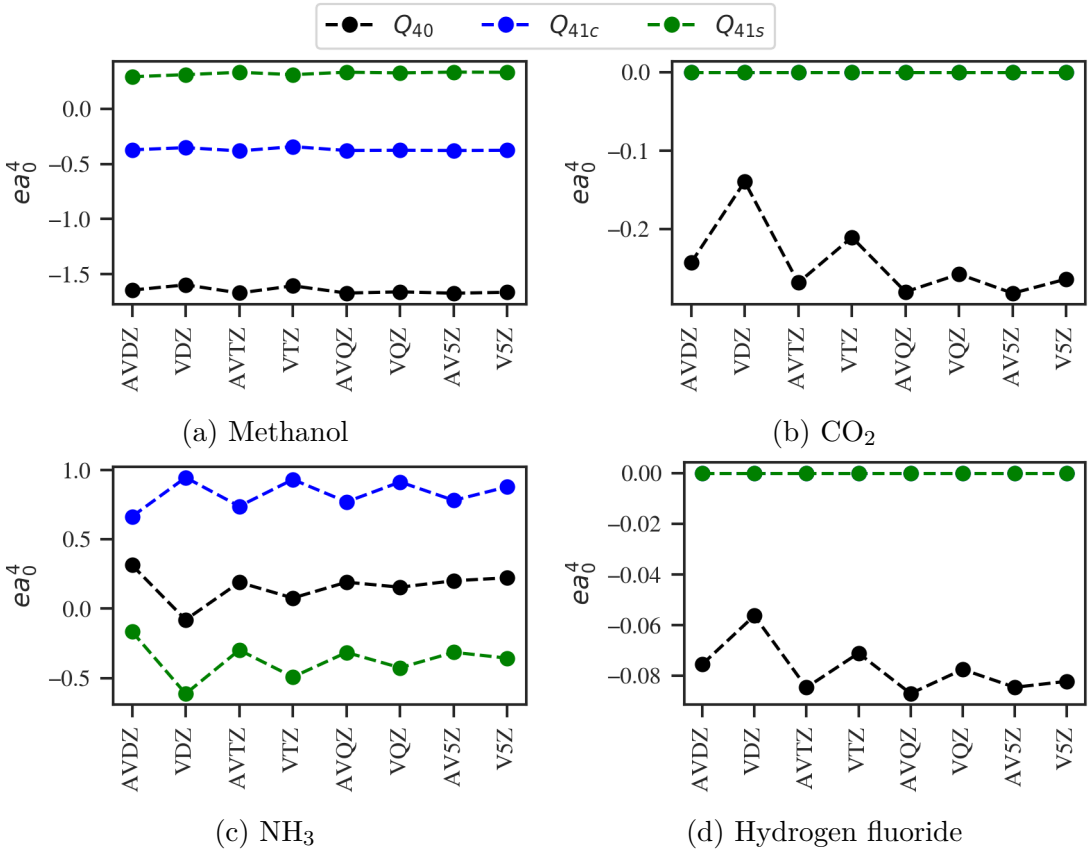
Figure C.6: Multipoles at 3rd order from SCF density matrix and GDMA.



(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

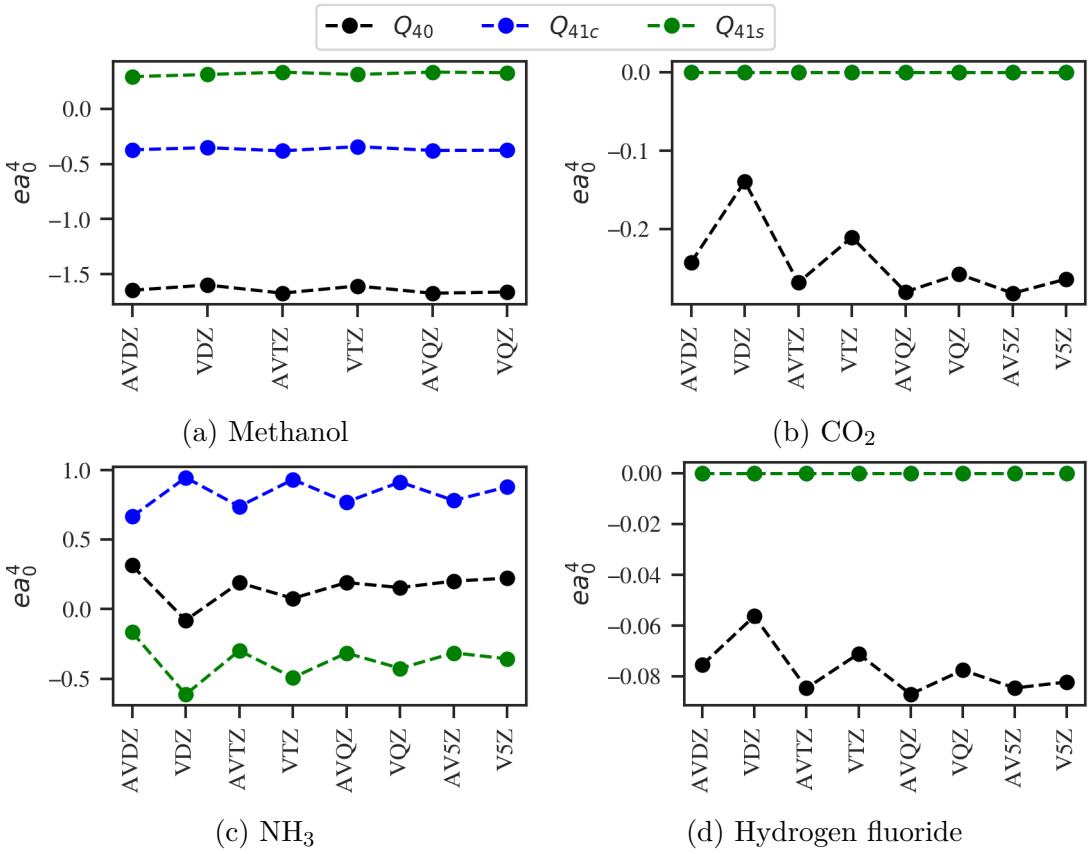
Figure C.7: Multipoles at 3rd order from CCSD(T) density matrix and GDMA.



(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

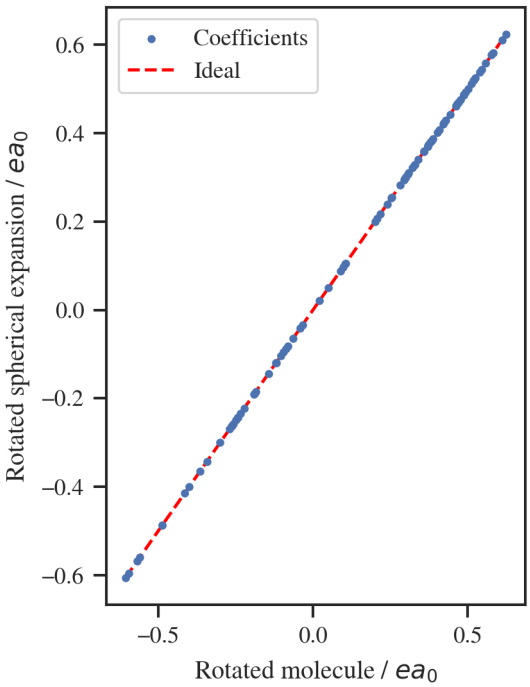
Figure C.8: Multipoles at 4$^{\text{th}}$ order from SCF density matrix and GDMA.



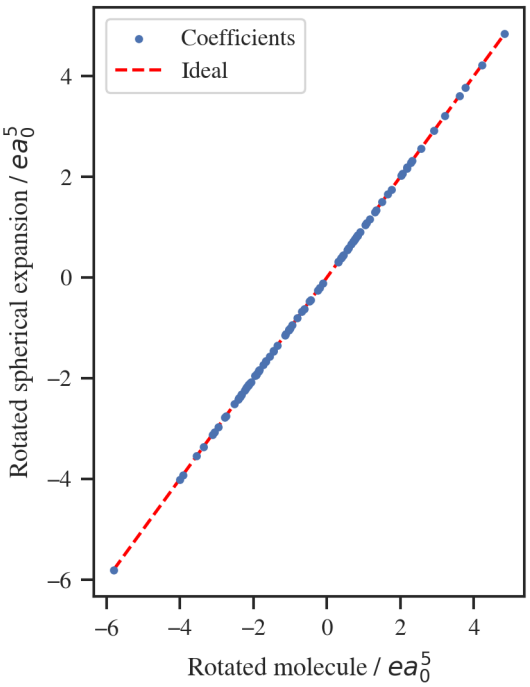
(a) Methanol (b) CO_2

(c) NH_3 (d) Hydrogen fluoride

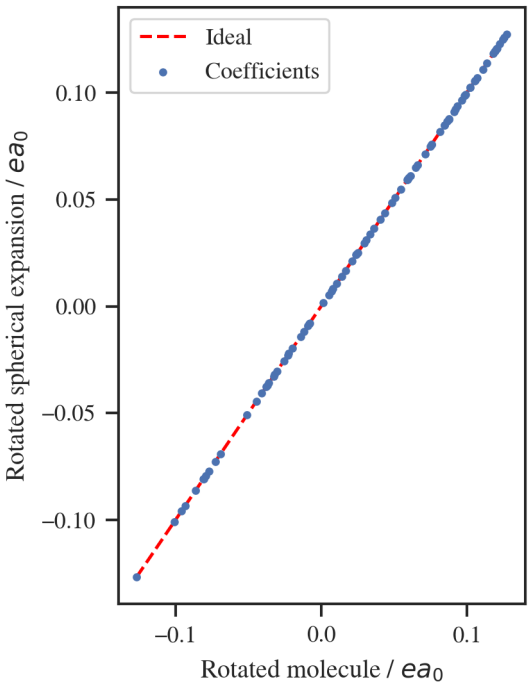
Figure C.9: Multipoles at 4$^{\text{th}}$ order from CCSD(T) density matrix and GDMA.



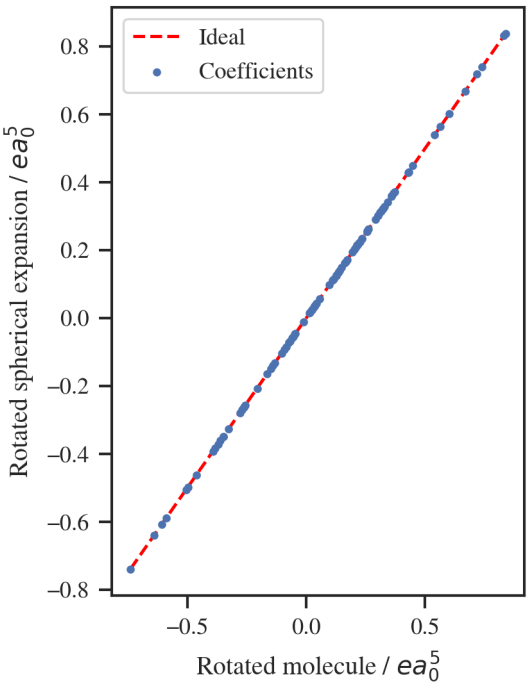
(a) Methanol C atom in Molpro DMA code and aug-cc-pV5Z in the first order of expansion.



(b) Methanol C atom in Molpro DMA code and aug-cc-pV5Z in the fifth order of expansion.



(c) Methanol C atom in Psi4 and GDMA code and aug-cc-pV5Z in the first order of expansion.



(d) Methanol C atom in Psi4 and GDMA code and aug-cc-pV5Z in the fifth order of expansion.

Figure C.10: Comparison of coefficients of multipoles obtained by rotation of multipole expansion in the spherical tensor formalism and by rotation of the density matrix and subsequent evaluation of the same coefficients.